DOCUMENT RESUME

ED 294 460 FL 017 396

AUTHOR Auger, Julie, Ed.
TITLE Actes du colloque: Tendances actuelles de la recherche sur la langue parlee (Conference Presentations: Current Trends in Research on Oral Language).
INSTITUTION Laval Univ., Quebec (Quebec). International Center for Research on Bilingualism.
REPORT NO CIRB-B-166; ISBN-2-89219-188-2
PUB DATE 88
NOTE 168p.; Papers presented at a student sociolinguistics conference at the University Laval (Quebec, Canada, September 25-27, 1987).
PUB TYPE Collected Works - Conference Proceedings (021)
LANGUAGE French; English

EDRS PRICE MF01/PC07 Plus Postage.
DESCRIPTORS Chinese; Discourse Analysis; Foreign Countries; *French; Grammar; Italian; Language Research; *Language Usage; Language Variation; *Linguistic Theory; Males; Multilingualism; *Oral Language; Phonetics; *Phonology; Regional Dialects; Research Methodology; Semantics; Sex Differences; Spanish; Verbs
IDENTIFIERS *Canada; *French (Canadian); Ontario (Hull); Ontario (Ottawa); Quebec; Quebec (Montreal); Quebec (Quebec)

ABSTRACT

Student papers presented at a sociolinguistics conference include: "La variation phonologique [d'une voyelle] a Hull"; "La force illocutoire des pronoms personnels 'je' et 'tu' en tant qu'insignes des places d'ou parlent les femmes et les hommes"; "L'emploi des modes indicatif et subjonctif dans le francais parle de la ville de Quebec"; "A Study of the (-ing) Variable in the Ottawa Male Community"; "An Application of Correspondence Analysis to the Multiple Variant Problem"; "Une etude des qualificateurs, intensificateurs et phrases encloses par rapport a la differenciation linguistique selon le sexe"; "Organisation conversationnelle en situation de plurilinguisme: les choix linguistiques des italophones de Montreal"; "'Pi' interrogatif en quebecois"; "Analyse parametrique de la chute du /s/ en espagnol"; "Lexical Variation in the Speech of Four Chinese Speakers"; "Marqueurs de discours et fonction discursive des narrations"; "Variation phonetique dans l'emploi des pronoms de troisieme personne en francais quebecois"; "L'alternance entre les determinants demonstratifs ['ce et cette'] en francais parle a Montreal"; and "A la recherche des diphtongues nasales du francais parle a Quebec". (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED294460

publication
B-166

centre international de recherche sur le bilinguisme

ACTES DU COLLOQUE

TENDANCES ACTUELLES DE LA RECHERCHE SUR LA LANGUE PARLÉE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
J. Pujiner
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

UNIVERSITÉ LAVAL
25 ET 26 SEPTEMBRE 1987

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✔ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Julie Auger (présentation)

1988

CIRB
ICRB

BEST COPY AVAILABLE

FL017396

Julie AUGER, présentation

Collaborateurs:
Richard Grenier - René Lapalme
Jean-François Montreuil - Paul Whitmore

ACTES DU COLLOQUE

TENDANCES ACTUELLES DE LA RECHERCHE SUR LA LANGUE PARLÉE

Publication B-166

1988
Centre international de recherche sur le bilinguisme
International Center for Research on Bilingualism
Québec

Le Centre international de recherche sur le bilinguisme est un organisme de recherche universitaire qui reçoit une contribution du Secrétariat d'État du Canada pour son programme de publication.

The International Center for Research on Bilingualism is a university research institution which receives a supporting grant from the Secretary of State of Canada for its publication programme.

© **1988 CENTRE INTERNATIONAL DE RECHERCHE SUR LE BILINGUISME**
Tous droits réservés. Imprimé au Canada
Dépôt légal (Québec) 1er trimestre 1988
ISBN 2-89219-188-2

TABLE DES MATIÈRES

Avant-propos i

Programme du colloque iii

La variation phonologique du [ɛ:] à Hull 1
Claire Pérusse

La force illocutoire des pronoms personnels *je* et *tu* en tant qu'insignes des places d'où parlent les femmes et les hommes 11
Sonia Morin

L'emploi des modes indicatif et subjonctif dans le français parlé de la ville de Québec . . . 27
Julie Auger

A study of the (-ing) variable in the Ottawa male community. 35
Dianne Fai

An application of correspondance analysis to the multiple variant problem 41
Wladyslaw Cichoki

Une étude des qualificateurs, intensificateurs et phrases encloses par rapport à la différenciation linguistique selon le sexe. 53
Charlene Potter

Organisation conversationnelle en situation de plurilinguisme: les choix linguistiques des italophones de Montréal 61
Normand Labrie

Pi interrogatif en québécois 75
Gilberte Léger

Analyse paramétrique de la chute du /s/ en espagnol 83
Valérie Catrice

Lexical variation in the speech of four Chinese speakers 99
Xu Daming

Marqueurs de discours et fonction discursive des narrations 121
Charleen Rains

Variation phonétique dans l'emploi des pronoms de troisième personne en français québécois 129
Marie-Josée Bourget

L'alternance entre les déterminants démonstratifs [Sə], [Sɛt] et [STə] en français parlé à Montréal 141
Michelle Daveluy

À la recherche des diphtongues nasales du français parlé à Québec 157
Sylvie Dubois

AVANT-PROPOS

Les 25 et 26 septembre 1987, se tenait à l'Université Laval un premier colloque-étudiant de sociolinguistique. Cette rencontre était organisée par l'Association des étudiants diplômés inscrits en langues et linguistique à l'Université Laval (A.É.D.I.L.L.). Le but était de rassembler les étudiants et les professeurs qui oeuvrent dans le domaine de la sociolinguistique afin qu'ils puissent échanger sur leurs recherches dans une ambiance le plus informelle possible. Nous espérions ainsi que des liens puissent se créer entre les chercheurs des différentes universités de l'Est du Canada.

Les résultats ont dépassé toutes nos attentes: dix-huit étudiants représentant six universités nous ont alors présenté leurs recherches en cours et les résultats de travaux récemment terminés. Les communications touchaient à peu près tous les domaines de la linguistique de même qu'une grande variété de langues. Nous avons délibérément choisi de ne pas limiter les communications à un seul type de sociolinguistique: un de nos objectifs était en effet de prendre connaissance des différents types de recherches qui s'offrent à tous ceux qui s'intéressent à la sociolinguistique. Environ 80 personnes ont participé au colloque et ont échangé et discuté avec les présentateurs.

Ce colloque nous a de plus donné l'occasion d'accueillir en nos murs, pour la première fois à Québec, le professeur William Labov de l'Université de Pennsylvanie, qui a accepté de prononcer une conférence sur "La recherche et la découverte de l'inconnu en sociolinguistique". M. Labov a semblé très heureux de constater à quel point la recherche sociolinguistique est dynamique dans nos universités.

Il va sans dire qu'un tel colloque ne pourrait se réaliser sans la participation d'un grand nombre de personnes. Il me faut d'abord mentionner que cet événement est vraiment le fruit d'un travail d'équipe et que sans l'implication et l'enthousiasme de Joyce Angio, le colloque n'aurait jamais eu lieu.

Je tiens ici à remercier les étudiants qui nous ont fait part de leurs recherches de même que tous ceux qui sont venus les écouter et qui ont participé aux discussions. Je voudrais aussi remercier M. Labov d'avoir accepté notre invitation car sa présence parmi nous fut très appréciée. Plusieurs membres de l'A.É.D.I.L.L. nous ont aussi grandement aidées lors de la préparation et de la tenue du colloque et c'est grâce à eux que tout s'est déroulé sans anicroche. Je remercie enfin Annette Dominik, Normand Labrie, Philippe Plamondon et Johane Verville d'avoir accepté d'agir en tant que présidents de séance.

Je tiens aussi à remercier toutes les instances de l'Université Laval qui nous ont appuyé financièrement pour l'organisation du colloque: le Département de langues et linguistique, la Faculté des lettres, dans le cadre des célébrations de son cinquantième anniversaire, l'Union des gradués inscrits à Laval (U.G.I.L.) et Vie étudiante. Je ne voudrais d'ailleurs pas passer sous silence le soutien moral du Département de langues et linguistique au cours de l'organisation et de la tenue du colloque.

Je tiens enfin à remercier de façon particulière Denise Deshaies, Claude Paradis et Diane Vincent, professeurs à l'Université Laval, qui nous ont communiqué leur enthousiasme pour la sociolinguistique. Leur appui et leurs bons conseils ont été grandement appréciés.

Ces actes présentent donc les textes de quatorze communications présentées lors du colloque. Certains de ces textes représentent les résultats de mémoires ou de thèses, d'autres présentent des recherches en cours alors que d'autres encore correspondent à des recherches exploratoires en vue

d'une étude plus approfondie; ils fournissent, je crois, une bonne idée des recherches menées par les chercheurs-étudiants dans les universités de l'Est du Canada. Leur publication a été rendue possible grâce à l'aide que m'ont apportée Richard Grenier, René Lapalme, Jean-François Montreuil et Paul Whitmore pour la révision et la correction des textes et grâce au Centre international de recherche sur le bilinguisme (C.I.R.B.) qui a accepté de les imprimer et de les diffuser.

L'expérience de ce type de rencontre entre étudiants oeuvrant dans le même domaine de la linguistique m'a semblé des plus positives et j'espère que ce n'est là que le début d'une série de colloques rassemblant les étudiants des différentes universités. Ce serait là une excellente façon de rompre l'isolement dans lequel nous travaillons, hélas, trop souvent.

Julie Auger
Ste-Foy

COLLOQUE

"TENDANCES ACTUELLES DE LA RECHERCHE SUR LA LANGUE PARLÉE"

PROGRAMME

VENDREDI, 25 septembre 1987

Local: Salon n^{o} 2 du Pavillon Maurice-Pollack

Président de séance: Normand Labrie

9h00 Accueil au salon n^{o} 1 du Pavillon Maurice-Pollack

9h30 Ouverture du colloque

9h45 **Claire Pérusse** (U. d'Ottawa)
"Exemple de variation phonologique"

10h15 **Sonia Morin** (U. de Sherbrooke)
"La force illocutoire des pronoms personnels JE et TU en tant qu'insigne des places d'où parlent les femmes et les hommes"

10h45 PAUSE

11h00 **Julie Auger** (U. Laval)
"L'emploi des modes indicatif et subjonctif dans le français parlé de la ville de Québec"

11h30 **Dianne Fai** (U. d'Ottawa)
"A Study of the (-ing) Variable in the Ottawa Male Community"

12h00 DINER: Buffet offert dans le cadre du colloque au local 3244 du Pavillon Charles-de-Koninck

Présidente de séance: Johane Verville

14h00 **Wladyslaw Cichoki** (U. du Nouveau-Brunswick)
"Application de l'analyse des correspondances à l'étude de la variation vocalique"

14h30 **Joyce Anglo** (U. Laval)
"Variation phonétique et réseaux sociaux: comportement langagier en milieu de contact des langues"

15h00 **Charlene Potter** (U. d'Ottawa)
"Study of Hedges, Qualifiers and Intensifiers in Relation to Linguistic Sex Differenciation"

15h30 PAUSE

15h45 **Normand Labrie** (U. Laval)
"Organisation conversationnelle en situation de plurilinguisme: les choix linguistiques des italophones de Montréal"

16h15 **William Labov** (U. of Pennsylvania)
"Recherche et découverte de l'inconnu en sociolinguistique"

17h15 Fin de la journée

■ ■ ■

SAMEDI, 26 septembre 1987

Local: Salon n^{o} 2 du Pavillon Maurice-Pollack

Présidente de séance: Annette Dominik

9h00 **Gilberte Léger** (U. de Sherbrooke)
"PI interrogatif en québécois"

9h30 **Sali Tagliamonte** (U. d'Ottawa)
"L'anglais de Samana: le temps du passé et un rapport préliminaire du -s verbal"

10h00 **Valérie Catrice** (U. du Québec à Montréal)
"Analyse paramétrique de la chute du /s/ en espagnol"

10h30 PAUSE

10h45 **Xu Daming** (U. d'Ottawa)
"Lexical Variation in the Speeches of Four Chinese Speakers"

11h15 **Charlene Rains** (U. de Montréal)
"Le rôle des marqueurs discursifs dans la structure narrative"

11h45 **Marie-Josée Bourget** (U. Laval)
"Variations phonétiques dans l'emploi de pronoms personnels de troisième personne en français québécois"

12h15 DINER

Président de séance: Philippe Plamondon

13h30 **Keltie Purcell** (U. d'Ottawa)
"L'insécurité linguistique dans une communauté acadienne de la Nouvelle-Écosse"

14h00 **Michelle Daveluy** (U. de Montréal)
"Changement dans la communauté francophone de Montréal: les déterminants démonstratifs en 1971 et en 1984"

14h30 **Sylvie Dubois** (U. Laval)
"La diphtongaison des voyelles nasales: approche dialectologique et sociolinguistique"

15h00 **Christopher Miller** (U. d'Ottawa)
"Contact linguistique: les conséquences linguistiques de l'emprunt lexical"

15h30 Mot de la fin

■ ■ ■

Nous tenons à remercier les organismes suivants de leur participation financière pour la tenue de ce colloque; sans eux, ce colloque n'aurait pu avoir lieu:

- *Département de langues et linguistique*
- *Faculté des lettres, dans le cadre des célébrations du 50e anniversaire de la Faculté*
- *Union des gradués inscrits à Laval (UGIL)*
- *Vie étudiante*

LA VARIATION PHONOLOGIQUE DU [ɛ:] À HULL[1]

Claire Pérusse
Université d'Ottawa

Les diverses recherches effectuées sur la "langue dans son contexte social" ont souligné la nécessité et la pertinence d'intégrer à l'analyse de facteurs d'influence linguistique l'effet de variables sociologiques pour cerner véritablement les liens qui existent entre la structure du langage et celle de son utilisation. A la lumière de ces réflexions, nous proposons de rendre compte d'un type de variation observée chez la communauté francophone de Hull, soit l'alternance phonologique de la variable (ɛ:) en variante de prestige [ɛ:] et en variante diphtonguée stigmatisée [aɛ], comme dans le mot "problème" (/prɔblɛ:m/ ~ /prɔblaɛm/).

Ce phénomène de diphtongaison semble remonter, selon certaines sources, à une époque assez lointaine et serait d'origine gallo-romane. Dans la préface de l'ouvrage de Marcel Juneau (1972) sur l'histoire de la prononciation française au Québec, Straka précise que certains traits de la prononciation québécoise se sont éteints à la suite d'une dédialectalisation progressive, alors que d'autres se sont maintenus jusqu'à nos jours, comme c'est le cas de la diphtongaison du [ɛ:]. Par l'analyse systématique des graphies attestées dans des documents anciens, Juneau (1972) constate déjà au XVIII[e] siècle la tendance actuelle de la diphtongaison de la voyelle longue [ɛ] par l'ouverture du segment initial (ɛ:>aɛ). Sur le plan articulatoire, la diphtongaison du [ɛ] a été décrite par Santerre (1971: 335): "par une remontée du dos de la langue vers le palais dur accompagnée d'un dégagement consécutif du passage pharyngal et de la région vélaire (...). Sur le plan acoustique, la diphtongue de [ɛ] est ouvrante."

L'objectif de la présente étude est de vérifier si la variable (ɛ:) et ses variantes de prestige [ɛ:] et diphtonguée [aɛ] sont en variation libre ou si le choix d'une variante sur une autre est régi par un ensemble de facteurs linguistiques et extra-linguistiques prévisibles. Les données ont été recueillies à l'intérieur du corpus du "Français parlé dans la région d'Ottawa-Hull"[2], constitué par Shana Poplack.

Parmi les différentes formes de différenciation sociale, notre analyse de la variation phonologique porte sur la stratification sociale, définie par Trudgill (1974a) comme une classification hiérarchique des individus à l'intérieur de la société. Il s'agit, en ce qui nous concerne, de la stratification à l'intérieur de deux classes sociales distinctes, composées d'individus présentant, pour chacune d'elles, les mêmes caractéristiques économiques et sociales.

Notre sous-échantillon d'informateurs a été sélectionné dans deux quartiers francophones de Hull, soit Vieux-Hull et Mont-Bleu. Au sein de la communauté francophone de Hull, ces deux quartiers se situent, chacun, à l'opposé de l'échelle socio-économique. Nous nous sommes intéressée à la variation chez deux sous-groupes particuliers: d'une part, les ouvriers non spécialisés du Vieux-Hull, comprenant trois hommes et trois femmes; et d'autre part, les travailleurs professionnels de Mont-Bleu, comprenant aussi trois hommes et trois femmes.

---

[1]Je tiens à remercier Shana Poplack pour les commentaires judicieux qu'elle a émis lors de la version préliminaire de cette analyse et dont a su profiter la présente étude. Ce travail s'inscrit dans le cadre d'un cours de maîtrise.

[2]cf. Shana Poplack (1985).

Dans une étude sur la stratification sociale du (r) à New York, Labov (1966) avait émis l'hypothèse générale que si deux sous-groupes de locuteurs sont classifiés selon un ordre quelconque dans la hiérarchie sociale, cet ordre transparaîtra aussi dans les différences entre ces deux groupes quant à l'utilisation de la variable étudiée. Nous avons tenté de vérifier si l'hypothèse énoncée par Labov se confirme également pour la variable (ɛ:) du français.

Le sexe constitue le second paramètre de différenciation sociale. Plusieurs études effectuées dans le but d'évaluer les disparités catégorielles et sociales véhiculées dans le langage (cf. Labov (1976), Fisher (1958), Trudgill (1972, 1974, 1983)) démontrent, entre autres, que les femmes - indépendamment d'autres variables telles que l'âge, l'éducation et la classe sociale - tendent à utiliser des formes linguistiques qui s'approchent davantage de la variante standard que les hommes en général. Trudgill précise, de plus, que ces différences de tendance entre le discours des hommes et des femmes reposent sur une question de degré et que ces discours se distinguent par une fréquence qualitative et quantitative des occurrences de la variable standard.

Outre la classe socio-économique et le sexe, nous avons tenté d'évaluer la pertinence des facteurs âge et niveau d'éducation, facteurs reproduits au Tableau 1. Cependant, pour satisfaire à nos deux principaux critères de sélection pour la classe socio-économique et le sexe, l'âge a été jugé secondaire. C'est ce qui explique l'absence de locuteurs très jeunes dans notre sous-échantillon de population. Les femmes dites "ménagères" ont été classées dans le corpus selon le statut socio-économique du mari ou du conjoint. De plus, pour garantir l'anonymat, chaque informateur a été pourvu, dans le corpus, d'un pseudonyme, puis d'un numéro qui apparaît ici entre parenthèses.

TABLEAU 1

VIEUX-HULL

	SEXE	AGE	ÉDUCATION	STATUT SOCIAL
(077)	M	40	primaire	éboueur (sur l'assurance-chômage)
(080)	M	46	primaire	homme à tout faire (bénéficiaire du bien-être social depuis 4 ans)
(081)	M	61	primaire	chauffeur de taxi (à la retraite)
(089)	F	41	primaire	"cleaning" (mari sur le "cleaning")
(090)	F	37	primaire	ménagère (mari éboueur)
(092)	F	52	primaire	ménage (mari?)

MONT-BLEU

	SEXE	AGE	ÉDUCATION	STATUT SOCIAL
(102)	M	35	post-secondaire	fonctionnaire
(106)	M	60	secondaire	dessinateur commercial (à la retraite)
(107)	M	70	post-secondaire	juge de paix (encore actif)
(113)	F	39	secondaire	ménagère (mari fonctionnaire)
(114)	F	36	secondaire	travaille dans une banque (mari gérant de banque)
(117)	F	62	secondaire	ménagère (mari fonctionnaire)

Pour chacun des informateurs, nous avons sélectionné cinquante occurrences de discours où l'on retrouve la variable (ɛ:), répartie de manière égale selon le style familier et le style plus ou moins formel. Il a été considéré comme relevant du style plus ou moins formel une conversation sur la langue, et comme familier, une conversation portant sur le quartier, l'école, le bon vieux temps, les histoires, etc.

Après une écoute plus attentive, nous avons dû rejeter certaines données de notre analyse, en conservant cependant cinquante occurrences par locuteur. A titre d'exemple, les déterminants ont été écartés du contexte variable, car bien qu'ils soient prononcés [ɛ:] en français standard, ils présentent chez les locuteurs de la région étudiée un phonème sous-jacent /e/ qui alterne avec [ɛ:] et non une forme sous-jacente /ɛ/ qui serait susceptible de donner lieu à une diphtongaison. C'est également le cas de la copule de troisième personne du singulier du temps présent.

À la lumière des études antérieures et suite à la prise en compte des facteurs d'influence linguistique et extra-linguistique sur la variation du [ɛ:]~[aɛ] auprès des locuteurs du Vieux-Hull et de Mont-Bleu, nous avons formulé les deux hypothèses suivantes:

1. La variation apparaîtra en position médiale de syllabe accentuée, entravée par une consonne allongeante ou devant une consonne non allongeante pour des raisons d'origine historique et étymologique (par exemple, l'accent circonflexe dans "fête" ou l'orthographe "ai" dans "maître" précédant une occlusive sourde jugée normalement abrégeante).

2. Les locuteurs de statut socio-économique inférieur et parmi eux, les ouvriers non spécialisés de sexe masculin seront plus enclins à utiliser la variante stigmatisée, [aɛ], que les autres groupes observés.

Dans 98% de nos données spécifiques, la variation apparaît sans équivoque en position médiale de syllabe. Quoique peu significatives, les occurrences figurant en positions initiale et finale de syllabe peuvent s'expliquer, en partie, par un style emphatique ou par des caractéristiques idiosyncrasiques qui témoignent d'une variation inhérente individuelle. Les données regroupées révèlent que les possibilités de retrouver la variante de prestige et la variante diphtonguée dans les mêmes contextes sont à peu près égales. On peut toutefois penser qu'une différence pourra vraisemblablement s'intensifier si l'on réduit le champ d'observation aux quartiers, aux locuteurs ou aux contextes lexicaux.

La répartition des occurrences entre les deux quartiers révèle au tableau 2 une nette tendance à la diphtongaison pour les locuteurs du Vieux-Hull, soit à 66,7%, contre 36% pour les locuteurs de Mont-Bleu. D'après ces résultats, les professionnels de Mont-Bleu prononcent la variante de prestige [:] dans les deux tiers des cas, et les ouvriers non spécialisés du Vieux-Hull ne l'emploient que dans 33,3% des cas. Mais ici, classe socio-économique et quartier ne sont pas des facteurs indépendants l'un de l'autre en raison de la constitution du sous-échantillon de population qui a été faite pour l'analyse.

TABLEAU 2

Répartition de la variation pour Vieux-Hull et Mont-Bleu

Variantes	Vieux-Hull	Mont-Bleu
[ɛ:]	33,3% (100)	64,0% (192)
[aɛ]	66,7% (200)	36,0% (108)
N	100% (300)	100% (300)

Pour l'étude de la variation, nous avons examiné le contexte phonologique post-vocalique pouvant exercer une influence sur la longueur de la voyelle, c'est-à-dire là où elle tend à s'allonger pour ensuite se diphtonguer (cf. Locke, 1949; Gendron, 1966; Jackson, 1969; Santerre, 1971; Juneau, 1972; Yaeger et Bourdeau, 1981). Au tableau 3, l'étude de l'influence des consonnes post-vocaliques pour l'ensemble des locuteurs indique que des quatre segments post-vocaliques retenus [r,m,t,z], l'occlusive est la seule consonne favorisant de loin la variante de prestige à 63%, alors que pour la nasale et la fricative, toutes deux sonores, la tendance est opposée.

TABLEAU 3

Influence des consonnes post-vocaliques pour l'ensemble des locuteurs

Variantes	/z/	/m/	/r/	/t/
[ɛ:]	34,7% (16)	39,7% (46)	50,7% (173)	63,0% (29)
[aɛ]	65,3% (30)	60,3% (70)	49,3% (168)	37,0% (17)
N	100% (46)	100% (116)	100% (341)	100% (46)

L'influence des segments consonantiques post-vocaliques se profile de manière distincte pour les ouvriers du Vieux-Hull et les professionnels de Mont-Bleu comme on peut le voir aux tableaux 4 et 5. Pour Vieux-Hull, si la fricative ne figure pas au tableau 4, c'est que nous n'avons conservé que les segments offrant un minimum de 20 occurrences.

TABLEAU 4

Influence des consonnes pour Vieux-Hull

Variantes	/m/	/r/	/t/
[ɛ:]	26,5% (13)	33,0% (61)	52,4% (11)
[aɛ]	73,5% (36)	67,0% (124)	47,6% (10)
N	100% (49)	100% (185)	100% (21)

TABLEAU 5

Influence des consonnes pour Mont-Bleu

Variantes	/z/	/m/	/r/	/t/
[ɛ:]	33,3% (13)	49,0% (33)	72,0% (112)	72,0% (18)
[aɛ]	66,7% (18)	51,0% (34)	28,0% (44)	28,0% (7)
N	100% (31)	100% (67)	100% (156)	100% (25)

Du fait que les locuteurs du Vieux-Hull utilisent davantage la variante diphtonguée, on pouvait s'attendre aussi à ce que ceci se réflète dans l'emploi respectif de la variable devant les segments consonantiques - ce qu'illustrent les tableaux 4 et 5. Ces résultats indiquent, de plus, que les segments [r,m,t,z] exercent une action différente sur la variation, selon les locuteurs et les quartiers considérés. Il faut en conclure soit que la contrainte phonologique n'est pas significative, soit que les pourcentages obtenus sont dus à la concentration dans l'un ou l'autre groupe de certaines unités lexicales qui affectent particulièrement la diphtongaison. Mais si l'on examine l'effet du segment illustré aux tableaux 4 et 5, en écartant les différences de fréquences absolues, on peut observer que dans les deux cas, la consonne nasale est plus favorable à la diphtongaison que la liquide et l'occlusive.

De manière parallèle, les consonnes voisées favorisent la diphtongaison dans une proportion supérieure aux consonnes sourdes. Les différences de fréquence entre les deux sous-groupes de population sont dues au fait, déjà mentionné, que le taux de diphtongaison est plus élevé au Vieux-Hull qu'à Mont-Bleu. La même courbe se dessine également au niveau de l'accentuation de la syllabe, tel que l'indique le tableau 6.

TABLEAU 6

Influence de l'accentuation sur la variation pour l'ensemble des locuteurs

Variantes	Accentué	Inaccentué
[ɛ:]	42,3% (211)	80,2% (81)
[aɛ]	57,7% (288)	19,8% (20)
N	100% (499)	100% (101)

En contexte accentué, la variante diphtonguée est prononcée dans 57,7% des cas. L'écart s'avère beaucoup plus perceptible en contexte inaccentué où dans 80,2% des cas la variable (ɛ:) figurera sous la réalisation de la variante longue. Un regard sur les quartiers aux tableaux 7 et 8 indique de plus que tous les locuteurs réalisent beaucoup plus la variante [ɛ:] en contexte inaccentué, dans une proportion de 67,3% au Vieux-Hull et de 91% à Mont-Bleu.

TABLEAU 7

Influence de l'accent pour Vieux-Hull

Variantes	Accentué	Inaccentué
[ɛ:]	27,0% (69)	67,3% (31)
[aɛ]	73,0% (185)	32,7% (15)
N	100% (254)	100% (46)

TABLEAU 8

Influence de l'accent pour Mont-Bleu

Variantes	Accentué	Inaccentué
[ɛ:]	58,0% (142)	91,0% (50)
[aɛ]	42,0% (103)	9,0% (5)
N	100% (245)	100% (55)

En outre, de toutes ces occurrences, on constate au tableau 6 que la grande majorité de celles-ci se retrouvent en contexte accentué dans plus de 80% (499) des cas et que pour ce contexte, les locuteurs de Vieux-Hull diphtonguent beaucoup plus (73%) que ceux de Mont-Bleu.

La répartition des occurrences selon les classes syntaxiques fait voir, au tableau 9, une prépondérance des unités lexicales monosyllabiques et bisyllabiques, composées en majorité de substantifs (49,3%); suivent la classe des verbes (18,8%), celle des adjectifs (15%) et celle des adverbes (14,7%) et dans une moindre mesure, la classe prépositionnelle (2.2%).

TABLEAU 9

Répartition de chacune des classes syntaxiques pour tous les locuteurs

Substantif	Verbe	Adjectif	Adverbe	Préposition
49,3% (296)	18,8% (113)	15,0% (90)	14,7% (88)	2,2% (13)
		100% (600)		

La répartition des variantes selon les classes syntaxiques pour l'ensemble des locuteurs souligne au tableau 10 une tendance plus perceptible à la diphtongaison pour la classe adjectivale que pour les autres classes. Cette tendance est encore plus distincte dans l'analyse individuelle des deux groupes observés - ce qui suit ici aussi la courbe initiale.

TABLEAU 10

Répartition des variantes selon les classes syntaxiques pour l'ensemble des locuteurs

Variantes	Adjectif	Préposition	Adverbe	Substantif	Verbe
[ɛ:]	32,2% (29)	46,1% (6)	48,9% (43)	50,3% (149)	57,5% (65)
[aɛ]	67,8% (61)	53,9% (7)	51,1% (45)	49,7% (147)	42,5% (48)
N	100% (90)	100% (13)	100% (88)	100% (296)	100% (113)

Lors de l'écoute individuelle des locuteurs, nous avons pu noter la fréquence élevée de certaines unités lexicales. Des vingt occurrences jugées les plus récursives, nous n'avons retenu que neuf d'entre elles offrant un minimum de 10 réalisations chacune. Elles figurent au tableau 11.

TABLEAU 11

Répartition des variantes selon les unités lexicales

	cher	affaire	problème	hiver	misère	même	faire	mère	père
[ɛ:]	10%	22,7%	26,3%	30,0%	38,5%	45%	56,4%	84,2%	85,2%
	(1)	(12)	(5)	(3)	(5)	(40)	(35)	(16)	(23)
[aɛ]	90%	77,3	73,7%	70,0%	61,5%	55%	43,6%	15,8%	14,8%
	(9)	(41)	(14)	(7)	(8)	(49)	(27)	(3)	(4)
N	100%	100%	100%	100%	100%	100%	100%	100%	100%
	(10)	(53)	(19)	(10)	(13)	(89)	(62)	(19)	(27)

Ce qui soulève le plus d'intérêt de part et d'autre, c'est la courbe opposée qui se dessine pour la réalisation de la variable dans le cas des vocables "affaire" et "cher". Dans 90,5% des occurrences de "affaire", c'est-à-dire 38 cas sur 42, sur un total de 53 occurrences pour l'ensemble des locuteurs, la variable sera diphtonguée par les ouvriers du Vieux-Hull. C'est ce qui explique au tableau 11 le pourcentage élevé de diphtongaison, soit 77,3%. Pour ce qui est de "cher", toutes les occurrences sont prononcées par le locuteurs du Vieux-Hull et 90% d'entre elles sont diphtonguées. On assiste ici à une différence presque qualitative entre les deux groupes sociaux. Il semble même que les locuteurs du Vieux-Hull aient "relexicalisé" le vocable "affaire" en [aɛ], ce qui porte à croire, dans ce cas, que l'on ait non pas le phonème sous-jacent /ɛ:/, mais le phonème /aɛ/ qui alterne avec [ɛ:]. Un questionnaire lexical pourrait révéler d'autres exemples similaires de "relexicalisation".

À l'instar de la plupart des études sociolinguistiques qui ont démontré une influence significative du style contextuel sur les contextes variables, les locuteurs de l'un et de l'autre groupe social ne semblent pas ici influencés par le style, réflétant la tendance générale. Ces résultats peuvent être faussés, en partie, par les données qui ne permettent pas vraiment une analyse des styles contextuels.

Dans notre analyse, la variable extra-linguistique "sexe" joue un rôle important de différenciation linguistique, illustrée aux tableaux 12 et 13.

TABLEAU 12

Répartition de la variable selon le sexe pour Vieux-Hull

Variantes	Hommes	Femmes
[ɛ:]	21,3%(32)	45,3% (68)
[aɛ]	78,7% (118)	54,7% (82)
N	100% (150)	100% (150)

TABLEAU 13

Répartition de la variable selon le sexe pour Mont-Bleu

Variantes	Hommes	Femmes
[ɛ:]	60,7% (91)	67,3% (101)
[aɛ]	39,3% (59)	32,7% (49)
N	100% (150)	100% (150)

On peut voir une distribution très inégale entre les deux groupes d'hommes et les deux groupes de femmes. Dans les deux quartiers, les hommes diphtonguent plus que les femmes, tendance qui est beaucoup plus prononcée au Vieux-Hull, où la variante [aɛ] est plus courante.

On peut s'interroger ici sur les raisons qui incitent les ouvriers de sexe masculin à une plus forte utilisation de la forme stigmatisée que les autres groupes observés. L'étude que nous avons faite sur la variation phonologique du (ɛ:) à Hull ne nous permet pas d'affirmer comme Trudgill pour l'anglais à Norwich (1983) que, pour les locuteurs mâles de la classe ouvrière, la forme stigmatisée non-standard est la forme de prestige à laquelle seraient associées des caractéristiques de masculinité et de robustesse (Labov (1966). Les trois ouvriers non spécialisés du Vieux-Hull représentent ici un échantillon de population trop restreint pour offrir des résultats concluants. Mais dans la conversation portant sur la langue, aucun d'eux ne reconnaît une quelconque forme de prestige au dialecte standard de Paris et le locuteur (077), qui se distingue par un taux de diphtongaison de 82%, s'offusquait même de devoir modifier sa manière de parler devant une personne de statut social plus élevé.

Pour vérifier si dans le cas de la variation du (ɛ:), il s'agit d'une variation stable, d'une variation en progrès ou, au contraire, s'il n'y a rien qui permet de conclure qu'il y a changement en cours, les locuteurs ont été divisés selon deux groupes d'âge: les moins de 45 ans et les 45 ans et plus.

TABLEAU 14

Répartition de la variable selon l'âge pour l'ensemble des locuteurs

Variantes	- de 45 ans	+ de 45 ans
[ɛ:]	49,0% (147)	48,3% (145)
[aɛ]	51,0% (153)	51,7% (155)
N	100% (300)	100% (300)

D'après nos données du tableau 14, rien n'indique que la variation que nous observons à Hull soit en progrès ou présente quelque tendance de progrès. On ne peut affirmer non plus que l'on soit en présence d'une variable stable comme dans le cas de la variable (ing) de l'anglais. Pour cela, il

aurait fallu observer la variable à plusieurs stades de son évolution. Les données de notre sous-échantillon ne permettent pas non plus de vérifier si la variation se généralise à l'intérieur de la communauté francophone à Hull.

En dernier lieu, nous avons voulu mesurer le niveau de scolarité et l'influence plausible qu'il exerce sur la variation (se reporter au tableau 1). Nous avons divisé les locuteurs en niveau primaire et en niveaux secondaire / post-secondaire. Ceci correspond en tous points à la division des locuteurs selon le quartier et la classe socio-économique. Ici, il est difficile de mesurer l'influence réelle qu'exerce le niveau de scolarité sur la variation, et de déterminer ce qui ressortit à la classe socio-économique et ce qui dépend vraiment du niveau de scolarité.

Il ne semble cependant pas dépourvu d'intérêt que, en dépit du fait que les femmes de Mont-Bleu ne présentent qu'un niveau secondaire d'instruction, tandis que deux hommes sur trois sont de niveau post-secondaire, les femmes utilisent la variante de prestige davantage que les hommes. On peut se demander, dans le cas présent, si la scolarité est un facteur plus influençable que la classe sociale ou le sexe. La réponse semble venir des femmes du Vieux-Hull, toutes de niveau primaire comme leurs vis-à-vis masculins, et qui diphtonguent moins que ces derniers, soit à 43,7% contre 59%.

Conclusion

Suite à notre analyse nous pouvons conclure 1. que la variable (ɛ:) et ses variantes ne sont pas en variation libre mais qu'elles sont motivées par des facteurs tant linguistiques qu'extra-linguistiques soulignant une régularité systématique; 2. que l'ordre selon lequel les locuteurs figurent dans la hiérarchie sociale se réfléchit également dans les différences qu'ils présentent pour l'utilisation de la variable (ɛ:); 3. que le discours des hommes et des femmes se distingue par une fréquence quantitative des occurrences des variantes; 4. que de plus, parmi la classe sociale inférieure, les locuteurs de sexe masculin favorisent davantage la réalisation de la variante stigmatisée; 5. que certaines classes syntaxiques entraînent plus que d'autres la réalisation diphtonguée de la variable.

Le facteur âge ne nous indique pas si la variation du (ɛ:) dans la communauté francophone à Hull est en état d'expansion, de stabilité ou en voie de disparition. D'après nos données, il n'apparaît pas non plus qu'un contexte est plus favorable à la variation qu'un autre. Plusieurs facteurs sont imbriqués les uns dans les autres et s'influencent mutuellement.

Enfin, nous pouvons aussi déduire de notre étude de la variation phonologique que, quel que soit le groupe social, il y a variation au sein de ce groupe et tous les locuteurs sont soumis aux mêmes contraintes - offrant des fréquences divergentes certes - mais tout à fait notables. En outre, notre analyse des contextes linguistiques démontre que, bien que, dans l'ensemble, les ouvriers non-spécialisés du Vieux-Hull emploient davantage la variante stigmatisée, ils ne sont pas insensibles à la variante de prestige, ce que seule une étude empirique peut vérifier.

Cette brève analyse ne constitue que l'ébauche d'une étude qui reste encore à approfondir et qui pourrait révéler, par exemple, à travers les groupes d'âges ou les unités lexicales jusqu'à quel point la variation phonologique du (ɛ:) est ancrée dans les habitudes langagières de l'ensemble de la communauté francophone de Hull.

BIBLIOGRAPHIE

FISHER, John (1958), "Social influences on the choice of a linguistic variant", dans *Word*, 14, pp. 47-56.

GENDRON, J.-D. (1966), *Tendances phonétiques du français parlé au Canada*, Les Presses de l'Université Laval, Québec.

JACKSON, W. (1969), "Étude du système vocalique du parler de Gravelbourg (Saskatchewan)", dans Pierre Léon (éd.), *Recherches sur la structure phonique du français canadien*, pp. 61-78.

JUNEAU, M. (1972), *Contribution à l'histoire de la prononciation française au Québec*, Les Presses de l'Université Laval, Québec.

KROCH, A. (1978), "Toward a theory of social dialect variation", dans *Language in society*, 7, pp. 17-36.

LABOV, William (1966), *The social stratification of English in New York City*, Center for applied linguistics, Washington, D.C.

LABOV, William (1976), *Sociolinguistique*, Les Éditions de Minuit, Paris.

LOCKE, W. (1949), *Pronunciation of the French spoken at Brunswick Maine*, American Society.

MILROY, Leslie (1980), *Language and social networks*, B. Blackwell, Oxford.

POPLACK, Shana (1985), "The care and handling of a mega-corpus: The Ottawa-Hull French project", dans F. Fasold et D. Shiffrin (éds.), *Language function and use*, Harcourt Brace Jovanovich, Orlando.

SANTERRE, L. (1971), *Les voyelles orales dans le français parlé à Montréal*, Montréal.

SANTERRE, L. et MILLO, J. (1978), "Diphtongization in Montreal French", dans D. Sankoff (dir.), *Linguistic variation: Models and methods*, Academic Press, New York, pp. 173-184.

TRUDGILL, Peter (1972), "Sex, covert prestige and linguistic change in urban British English", dans *Language in society*, 1, pp. 179-195.

TRUDGILL, Peter (1974), *Sociolinguistics: An introduction*, England.

TRUDGILL, Peter (1974a), *The social differenciation of English in Norwich*, Cambridge University Press, Cambridge.

TRUDGILL, Peter (1983), *On dialect*, B. Blackwell, Oxford.

WALKER, D. (1984), *The pronunciation of Canadian French*, University of Ottawa Press, Ottawa.

YAEGER, M. et BOURDEAU, M. (1981), *Montreal French vowel durations*, Montréal.

LA FORCE ILLOCUTOIRE DES PRONOMS PERSONNELS *JE* ET *TU* EN TANT QU'INSIGNES DES PLACES D'OÙ PARLENT LES FEMMES ET LES HOMMES

Sonia Morin
Université de Sherbrooke

Introduction

Dans le présent article, il sera question des valeurs des pronoms personnels JE et TU dans le discours d'informatrices et informateurs du corpus des "Enquêtes sociolinguistiques dans la région de Sherbrooke". Les résultats d'une étude statistique (Beauchemin, 1983) effectuée sur ce corpus a révélé que le discours des femmes se distinguait de celui des hommes par une haute fréquence du pronom JE alors que celui des hommes était marqué par un emploi caractéristique du pronom TU. Ces résultats m'ont intriguée au point de chercher à travers un mémoire de maîtrise le pourquoi d'un tel comportement différentiel en matière de référence personnelle selon que l'on est femme ou homme. Ce mémoire a été dirigé par les professeurs Michel Théoret, études françaises, et Claude Nélisse, service social, de l'Université de Sherbrooke.

Il ne s'agit pas ici d'une démarche scientifique au sens de "faire la preuve de", ou encore de confirmer ou infirmer une théorie quelconque, ni d'ajouter une contribution à un modèle linguistique déjà existant, et encore moins de participer à une formalisation de la parole. Non! Ce qui est en jeu ici est un désir de comprendre au mieux ce qui se passe quand deux personnes se parlent. Et comme la compréhension d'un phénomène n'est jamais que le fruit d'un travail intellectuel qui prend ses assises dans l'approfondissement des nombreuses causes possibles de son avènement, la force d'une telle démarche réside dans la plausibilité des interrelations établies entre les divers champs d'investigation porteurs d'éléments productifs de faits de parole.

Ma recherche m'a amenée à dépasser le cadre de la linguistique dite traditionnelle parce qu'elle s'obstine à considérer les divers faits de parole comme des phénomènes purement linguistiques et à évacuer systématiquement l'extra-linguistique. Or je crois non seulement que l'extra-linguistique influence la production de discours mais j'ai également la certitude que la parole ne se réduit pas à la simple production d'énoncés dans le but de communiquer; elle est d'abord et avant tout essentielle à la constitution du sujet et à sa réalisation. C'est pourquoi je me suis aventurée dans les domaines de l'analyse interactionnelle et de la psycho-sociologie.

L'objet de recherche

L'intérêt d'une étude sur les pronoms personnels JE et TU repose sur le fait que ces deux unités font appel à la référence personnelle et à la reconnaissance de l'autre. En effet, JE et TU sont les seuls pronoms personnels qui posent clairement les questions **qui parle?** et **à qui JE parle?** Et il y a plus: JE n'existe qu'en sa qualité de sujet de l'énonciation et de l'énoncé, qualité que TU a le pouvoir de lui reconnaître ou pas. D'autre part, TU doit sa propre existence à JE puisque c'est lui qui l'interpelle. On le sent bien, ces deux pronoms sont indissociables et ils participent autant de la subjectivité du langage que de l'intersubjectivité qui unit les deux interlocuteur/trice/s.

L'échantillon sur lequel porte ma recherche a été tiré du corpus des "Recherches sociolinguistiques dans la région de Sherbrooke" (1972), lesquelles regroupent 49 informateurs et 51 informatrices, tous et toutes originaires de l'Estrie. Il se compose de 24 informateur/trice/s dont les paramètres sociologiques tels la nature du travail, l'âge et la scolarité ont été neutralisés de façon à ce que seul le sexe permette de différencier le premier groupe du second. Ainsi, le profil moyen du groupe féminin correspond à une femme travailleuse spécialisée, âgée de 39,3 ans et possédant une scolarité de 10,2 années. Du côté masculin, le profil type est celui d'un homme travailleur spécialisé, âgé de 39,4 ans et possédant 10,5 années de scolarité. Les 24 informateur/trice/s ont été interrogé/e/s par quatre enquêteur/trice/s: deux femmes, Lise et Gaétanne, et deux hommes, Paul et Guy, qui se sont présenté/e/s comme des étudiant/e/s travaillant pour une recherche universitaire.

L'étude statistique dont il a été question plus haut repose sur l'ensemble du corpus, c'est-à-dire sur les réponses aux 16 thèmes libres. Or, comme je désirais analyser en détail les interactions verbales, je n'ai retenu que le premier thème, jugé représentatif parce qu'il suscitait l'évaluation, l'auto-évaluation, la justification et l'opinion de l'informateur/trice, ce qui lui permettait de s'impliquer, s'il/elle le désirait, dans les réponses qu'il/elle fournissait.

Le décompte des pronoms JE et TU dans les réponses des 24 informateur/trice/s au premier thème a révélé que la variation tenait toujours, ce qui a confirmé que le thème choisi était bel et bien représentatif.

TABLEAU I

FEMMES		HOMMES	
JE	TU	JE	TU
160	25	90	45

À ce point-ci de l'étude, c'est-à-dire sans aucune analyse en profondeur et hors de tout contexte, il est permis de dire que les femmes ont utilisé 1,8 fois plus de JE que les hommes pour répondre aux questions du premier thème. De même, les hommes ont utilisé 1,8 fois plus de TU que les femmes dans leurs réponses aux mêmes questions.

En linguistique

Toutes les grammaires traditionnelles présentent JE comme la personne qui parle et TU comme la personne à qui l'on parle. Sans être fausse, cette affirmation se limite aux référents les plus fréquents des pronoms personnels JE et TU. Or JE et TU sont beaucoup plus complexes qu'ils ne le laissent supposer à première vue. Fondamentalement, il est vrai que ces deux pronoms constituent des déictiques purs, car si virtuellement ils peuvent représenter n'importe qui, actualisés, ils ne peuvent que référer aux deux personnes dans l'ici et maintenant de la prise de parole. C'est en approfondissant le contexte d'apparition qu'il est possible de voir les extensions de chacun de ces pronoms.

Le JE

Dans le cas de la première personne, il est impossible que JE ne puisse pas ne pas parler de lui-même. Toutefois, il peut le faire d'au moins deux façons: en étant celui qui dit, locuteur/énoncé (dictum), et celui qui critique ce qui est dit ou ce qu'il dit de lui-même, locuteur/énonciation (modus). C'est par l'étude des verbes auxquels s'accole JE qu'il est plus facile de percevoir les deux valeurs du pronom de la première personne. En effet, dans le dictum, on retrouve des assertifs simples, c'est-à-dire des verbes qui ne font que véhiculer une certaine quantité d'information. Dans le modus les verbes relèvent du regard que le locuteur pose sur l'énonciation et traduisent le procès que ce dernier fait de ce qu'il dit. Ce sont des assertifs complexes. Ce sont des verbes d'appréhension intellectuelle, d'appréhension perceptive, de volonté et de sentiment. Les verbes considérés comme semi-auxiliaires, modificateurs du mode et de l'aspect de l'infinitif, font aussi partie du modus. Ces verbes forment un groupe en raison d'un trait qu'ils détiennent en commun: ils régissent une deuxième assertion, proposition complétive, percontative ou infinitive. Ils peuvent prendre également le nom de modélisateurs. Par exemple, à la question **Considérez-vous que vous, vous parlez bien?**, un informateur a répondu

(1) "Hum. Non. Non. JE parle pas très bien." H104-22-26[1]

alors qu'une informatrice a dit

(2) "Non. Ah non! (...) JE peux pas dire que JE parle mal." F119-29-28

Dans le premier exemple, PARLER est un assertif simple qui dénote un comportement verbal alors que dans le second, POUVOIR est un assertif complexe qui modifie la valeur première de DIRE, lui-même assertif complexe qui régit une proposition complétive.

Le tableau suivant présente la liste des verbes les plus fréquemment employés avec le JE dans le discours des informatrices et, par comparaison, dans celui des informateurs.

TABLEAU II

	FEMMES	HOMMES
SAVOIR	23	12
PENSER	15	5
PARLER	13	4
POUVOIR	9	9
DIRE	9	8
ALLER	9	3
CROIRE	7	4
VOIR	7	3
TROUVER	6	6
FAIRE	4	1
VOULOIR	4	1
TOTAL	106	56

[1]H104-22-26 constitue la référence du texte, il renvoit à homme, n° d'enquête, page et ligne. Il en va de même pour F119-29-28, F renvoyant à femme.

Comme on peut le constater, la majorité des verbes du tableau II sont des modélisateurs, ils appartiennent donc au modus.

Le TU

Bien que TU, de par sa valeur première d'allocuteur, renvoie habituellement à la personne à qui l'on parle, il peut également faire référence à un individu ou à un groupe d'individus non-défini comme dans l'exemple suivant:

(3) "Quand T'es habitué de parler français, o.k. Si TU... Si TU parlais... TU travaillerais avec des Anglais là, là o.k. Ça serait un charme encore. Mais quand T'es habitué dans le français là, c'est... assez dur." F244-234-12

En remplaçant tous les TU par des ON, il est impossible de déterminer le référent exact d'un ou l'autre des pronoms. Par contre, on sent que l'équivalence n'est pas parfaite. Les deux énoncés ont une valeur de généralité mais cette généralité se distribue différemment selon le pronom employé. Avec ON elle s'applique pour tous; avec TU elle s'applique également pour tous, mais à raison d'un seul à la fois. C'est ce que Denise Deshaies appelle un "destinataire distributif" (Deshaies, 1982: 4). Autrement dit, l'emploi du TU non-défini a pour " effet de poser l'allocutaire comme référent hypothétique..." (Laberge, 1977: 97)

TU connaît également d'autres emplois "déviants". C'est le cas des expressions SI TU VEUX et TU VAS PRENDRE.

(4) "Ben, c'est-à-dire, j'haïs pas Jean Malo, SI TU VEUX, le midi là. Jean Malo, y va te dire son affaire là lui pis, bien parler, SI TU VEUX, comprenable. Tandis que TU VAS PRENDRE Jacques Tremblay, ça va être comme des tiges qu'y va mélanger avec qu'est-ce qu'y veut dire..." H103-1-20-/1-24

Ces deux expressions constituent des parenthétiques ou des incises dans lesquelles la structure sémantique relève de l'hypothèse, ou de la supposition. On peut décoder le référent comme étant n'importe qui, l'allocuteur/trice y compris/e.

Enfin, il y a le fameux T'SÉ, expression fort bien connue en français québécois, dont la valeur globale en est une d'allocution indirecte qui l'apparente à la locution **n'est-ce-pas**. L'étude du contexte large a permis d'identifier trois contextes d'apparition qui définissent la valeur de cette expression:

1) simple appuyeur de propos sans demande de sanction:

(5) "Ça y en a pas qui parlent parfaitement, parce que le parfait ça, c'est assez dur à obtenir. Mais, disons qu'on a été dans un milieu, T'SÉ, pour, pour être dans moyenne disons. On est pas pire que les autres." H112-30-1/30-4

2) demande de sanction (le texte entre parenthèses représente les éléments prosodiques):

(6) "Tandis que tu vas prendre Jacques Tremblay, ça va être comme des tiges qu'y va mélanger avec qu'est-ce qu'y veut dire... (baisse de l'intonation, E dit "oui,oui" mais ne reprend pas la parole) Ou encore ben y essaie à se faire des farces pour rire de lui-même... (baisse de l'intonation et long silence, E ne reprend pas ni ne sanctionne) T'SÉ, euh." (E rit un peu et reprend son tour de parole en disant "je comprends, je comprends" et ce faisant elle sanctionne la réponse de I.) H103-1-22/1-27

3) embrayeur pour une opinion personnelle exprimée par un JE ou par un "pour moi":

(7) "? Bon, d'accord. Pis êtes-vous quand même en faveur du bilinguisme?

Ben là là, T'SÉ, chus en faveur parce que pour gagner ta vie aujourd'hui, y te le faut..." H136-57-22/57-25

Il est maintenant temps de présenter la répartition des TU employés dans le discours des hommes et, par comparaison, dans celui des femmes.

TABLEAU III

HOMMES		FEMMES	
TU d'allocution	3	TU d'allocution	1
TU non-défini	11	TU non défini	18
TU (expression)	10	TU (expression)	-
T'SÉ	21	T'SÉ	6
TOTAL	45	TOTAL	25

Ainsi, le TU qui caractérise le discours des informateurs est celui des expressions SI TU VEUX, TU VAS PRENDRE et T'SÉ.

Si l'investigation en linguistique a permis de cerner quelles étaient les valeurs distinctives des pronoms personnels JE et TU qui marquent le discours des informatrices, JE+ modélisateur, et celui des informateurs, expressions idiomatiques et figées, elle n'a pas apporté d'explication quant au comportement différentiel des femmes et des hommes en ce qui a trait à l'emploi marqué d'un pronom plus qu'un autre.

L'analyse de la dynamique d'entrevue

En analysant la dynamique verbale particulière qu'est l'entrevue, il sera principalement question des stratégies des réponses et des questions en termes de pouvoir des enquêteur/trice/s et des informateur/trice/s dans la négociation de l'information.

L'entrevue

L'entrevue est une demande formelle d'information dans laquelle la parole est distribuée contractuellement en raison de l'assignation des rôles d'intervieweur/euse et d'interviewé/e. L'entrevue n'existe que si les participant/e/s ont accepté de jouer ces rôles. Il y a donc d'un côté l'intervieweur/euse, investi/e du pouvoir d'administrer le questionnaire, qui décide du début et de la fin de l'entrevue, qui en oriente le déroulement et qui sanctionne la pertinence des réponses en fonction de l'information qu'il/elle désire obtenir. C'est lui/elle qui pose des questions. De l'autre côté se trouve l'interviewé/e, détenteur/trice de l'information désirée, supposément, vu l'acceptation du contrat, désireux/euse de bien servir les fins de l'enquête. Il/elle répond aux questions. Alors qu'il semble à première vue que l'entrevue constitue une instance de discours sans surprise où tout devrait se dérouler rituellement, il est nécessaire de se rappeler qu'

> "une conversation, même si elle est prédéterminée par un certain nombre de facteurs socio-culturels, [notre exemple d'interview est ici significatif] n'est pas une réalité totalement pré-construite; elle se construit au fur et à mesure, et pour nombre de ses aspects (longueur, nombre d'interventions...) l'on peut dire que "ce n'est pas joué d'avance"." (Simonin, 1977: 23)

La question

Si la question constitue fondamentalement une demande d'information, elle est également

> "une prise effectuée sur un autre sujet parlant en le constituant, quoi qu'il fasse, en répondeur virtuel (...) Elle est une main-mise d'ordre symbolique sur le corps, le temps et la parole de l'autre du simple fait qu'elle brise le silence et ouvre un espace verbal: un espace d'échange langagier qui constitue par lui-même, parce qu'il est situé dans le champ général des rapports sociaux, un domaine dont l'appropriation et le contrôle sont immédiatement en jeu. Prise redoublée par le fait que la question oriente l'espace verbal qu'elle instaure ou, si l'on préfère, tente de contrôler la réponse." (de Fornel et Encrevé, 1983: 5-6)

Y a-t-il eu des facteurs reliés à la façon de poser les questions aux informateur/trice/s qui auraient pu influencer la production de la variation?

Dans le cas de la haute fréquence du JE+ modélisateur dans le discours des femmes, j'ai vérifié, dans un premier temps, si la modélisation des réponses avait été entraînée par une plus grande modélisation dans les questions adressées aux femmes que dans celles adressées aux hommes. Non! Un nombre semblable de questions comportant un modélisateur a été adressé aux informatrices et aux informateurs: 50 pour les femmes et 49 pour les hommes. Ce n'est donc pas un facteur qui a influencé l'emploi distinctif du JE+ modélisateur chez les informatrices.

Dans un second temps, je me suis demandé si les questions posées aux femmes contenaient une forte insistance dans l'adresse: l'emploi répétitif du pronom d'allocution (**vous ou tu**), que ce soit sous forme d'interpellation directe, ou encore de sollicitation d'opinions fondées sur l'expérience personnelle. Dans l'extrait d'entrevue suivant, Gaétanne a utilisé trois formes d'insistance d'adresse:

> "? Hum, hum. Oui. O.K. D'après toi, est-ce que toi-même, tu parles bien?
> F211-177-14/177-15

Encore une fois, la réponse est non. Mais il y a plus. Le nombre d'insistances est plus fort pour les hommes que pour les femmes: 97 insistances contre 71.

Nous devons reconnaître que les questions des enquêteur/trice/s ne comportent aucun facteur déterminant, du moins dans l'énoncé, qui aurait provoqué un emploi distinctif du JE+ modélisateur chez les femmes.

Dans le cas des expressions SI TU VEUX, TU VAS PRENDRE et T'SÉ, il semble impossible de trouver dans la dynamique de la question un facteur explicatif de leur haute fréquence dans le discours des hommes et ce, en raison même du fait qu'elles ont toutes une valeur d'allocution indirecte.

Dans un cas comme dans l'autre, je ne minimise pas l'influence des intervieweur/euse/s, mais je crois que l'explication du comportement différentiel des femmes et des hommes ne se trouve pas dans des facteurs "visibles" des questions qu'ils/elles ont posées mais plutôt dans des attitudes et des images qu'ils/elles ont dégagées, que les interviewé/e/s ont perçues et auxquelles ils/elles ont repondu.

La réponse

La réponse se conçoit sous deux aspects: information faisant suite à une demande d'information formulée par une question et réaction face à la prise effectuée sur le corps, le temps et la parole de l'interviewé/e. Son pouvoir réside dans les nombreuses stratégies que l'informateur/trice peut déployer pour atténuer, voire, dans certains cas, annuler, l'emprise que le/la questionneur/euse établit sur lui/elle. Toutes les stratégies interactionnelles sont issues de la résistance à cette emprise. La résistance se manifeste principalement par la négociation sur l'information. La réponse peut être littérale, partielle, contradictoire au lieu d'être sincère, complète et satisfaisante. L'hésitation, les questions en retour pour des précisions, les redondances et les digressions sont autant de stratégies pour se ménager du temps et imposer son propre rythme à l'entrevue. D'autres stratégies comme l'abandon de son propre discours au profit de celui de l'autre ou de celui du cadre dialogique strict, le refus explicite et l'hostilité ouverte constituent des attitudes radicales.

Puisque l'étude des questions de l'entrevue n'a révélé aucune explication satisfaisante pour la présente variation, j'ai concentré mon attention sur les réponses. Il semble bien que la seule certitude qui se dégage à ce stade-ci de l'étude est qu'effectivement la haute fréquence du JE du modus dans le discours des informatrices et l'emploi distinctif des expressions avec TU chez les informateurs, correspondent à des stratégies de la réponse dans laquelle le sujet inscrit son rapport à la parole.

Selon de Fornel et Encrevé, l'emploi de JE+ modélisateur constitue une attitude propositionnelle envers l'objet de la question qui inscrit dans le discours la possibilité d'une reprise, d'un retour sur ce qui a été dit (...). (de Fornel et Encrevé, 1983: 25) Quant aux expressions SI TU VEUX, TU VAS PRENDRE et T'SÉ, elles relèvent toutes d'une dynamique de sollicitation indirecte dans laquelle l'autre est considéré comme un destinataire distributif. Sollicitation indirecte parce que dans la majorité des cas, aucune sanction du propos n'est apportée par l'enquêteur/trice.

L'analyse de la dynamique d'entrevue a mis en lumière que c'est dans les réponses, du fait qu'elles contiennent le rapport que chacun/e entretient face à la parole comme insigne de place d'où il/elle parle, qu'il faut chercher les éléments explicatifs de la variation dans l'utilisation des pronoms personnels selon que l'on est femme ou homme.

Ce qu'est la parole

Pour F. Flahault, la parole ne saurait être réduite à un simple moyen de communication ou d'information, car s'il est vrai que l'énoncé donne quelque chose à connaître, le propos, il donne également quelque chose à reconnaître au-delà du visible qui tient du sujet et de sa double demande, demande de pertinence de son propos et demande de reconnaissance de la place d'où il parle. En fait,

> "la parole est mixte, intermédiaire entre ces deux registres: si captive qu'elle soit du discours qui la détermine et de la référence à un quelque chose dont il est parlé, elle ne cesse pour autant d'opérer dans l'ordre du "qui je suis pour toi, qui tu es pour moi"." (Flahault, 1978)

Le sujet

La parole ne saurait se comprendre sans aborder la délicate, pour ne pas dire épineuse, question de la constitution du sujet. La notion de sujet ne doit pas être confondue avec celle du moi. En effet, le moi correspond à un ensemble de perceptions et de représentations que l'individu se fait de sa propre image, alors que le sujet est cet être qui existe, qui pense et qui demande constamment qu'on reconnaisse son existence. Le premier est sémiologique, le second est réel.

La relation d'un sujet à un autre ne pourrait avoir lieu sans un ensemble de dispositions, de médiations, qui régularisent et tempèrent cette relation et qui permettent au(x) sujet(s) de s'identifier de façon limitée et spécifiée tout en lui(leur) fournissant une visée à un monde complet, cohérent et sans lacune, la complétude.

La parole régularise et tempère la relation d'un sujet avec un autre en offrant aux interlocuteurs un espace de réalisation. À la fois constituante du sujet et source des marques de reconnaissance, elle permet l'interaction subjective. En prenant la parole, le sujet

> "se réalise en tant que reconnu à une place qui elle-même se définit dans un système de places. Se réalise et non s'exprime puisque est en cause l'existence même du sujet et non pas la simple question de savoir comment il parle de lui. Place et non rôle, car c'est l'identité même du sujet que sa position dans un système de places soutient." (Flahault, 1978: 70).

La parole, en tant que lieu de constitution du sujet, garantit pour chacun le maintien de sa place fondamentale de parlant, peu importe l'assignation de place dont il est l'objet. Elle tend à produire des rapports de places pour contrer l'antagonisme que véhicule le concept de place: d'un côté, une place se définit par rapport à d'autres places, de l'autre, elle vise la complétude. Le rapport de places ne peut "être qu'un rapport où, tout en s'établissant comme limites, les places satisfont au mieux la visée de complétude." (Flahault, 1978: 165) Le concept de place renvoie toujours à deux choses: à la réalité matérielle, déterminée par la division du travail ou du sexe, et à une représentation qui vaut comme réalité puisque le sujet se réalise dans la parole. La place se définit par "les déterminations qui font que le système dont elle relève s'impose comme réalité et (...) -par- l'investissement du désir venu s'inscrire et se méconnaître dans tel rapport de places." (Flahault, 1978: 138) La spécificité du concept de place tient essentiellement au fait que chacun n'accède à son identité qu'à partir et qu'à l'intérieur d'un systéme de places qui le dépasse. En effet,

> "chaque sujet n'existe que s'il est reconnu qu'il existe, et il ne s'attire cette reconnaissance que s'il en produit le signe attendu. Produire au sens de manifester; car le sujet n'est jamais le créateur individuel de ce signe - sans quoi, évidemment, personne ne pourrait le reconnaître à ce signe (qui d'ailleurs n'en serait pas un)." (Flahault, 1978: 58)

Or, comme la reconnaissance n'est jamais l'objet explicite de l'interaction verbale, c'est sous le couvert de l'objet dont il est question dans la situation d'échange verbal qu'elle se traite et elle est alors implicite. Et c'est en raison du fait que "les individus ne sont pas maîtres d'opérer leur mise en place puisque c'est au contraire cette mise en place qui établit leur identité" (Flahault, 1978: 52) qu'il en est ainsi.

Les places d'où parle le sujet

Chaque fois que le sujet prend la parole, il le fait à partir de quatre registres de places qui lui permettent de s'identifier et de viser la complétude:

1) Le registre inconscient est celui des paroles et des relations qui ont structuré inconsciemment le sujet et qui ont fixé le cadre de ses relations transférentielles.

2) Le registre idéologique prend assise sur des systèmes discursifs qui correspondent à la place du sujet dans la formation sociale à laquelle il appartient. C'est le lieu où la division des places et des classes, inhérente aux rapports de production, entre en relation avec la vocation de tout discours à la complétude. Ici, la parole du sujet est issue de la place qu'il occupe au sein de la formation sociale. Comme l'individu se définit, en termes d'identité, à travers les représentations sémiologiques de la réalité, son discours s'inscrit dans une visée de complétude. Il tend à vivre sa place dans ce monde de représentations et sa parole se veut vraie, pertinente, totalisante et témoignant de sa haute vision des choses. Ce qu'il désire le plus est que l'autre reconnaisse les insignes dont il est porteur. Mais plus encore, s'il acquiert cette reconnaissance, c'est qu'il aura prouvé, par son discours, qu'il occupe bien la place qu'il croit occuper, ce qui aura pour effet de confirmer qu'il est bien dans la réalité.

3) Le registre institué par telle situation de parole est celui qui assigne une place au locuteur à l'intérieur du système de places qui rend possible la présente situation de parole. C'est le registre le plus facile à identifier, car il s'agit de savoir à quel titre la parole de chacun se produit.

4) Le registre de la circulation d'insignes dans le tissu discursif est celui dans lequel la place du sujet est en rapport avec le discours dominant dans la situation d'interaction verbale. C'est le lieu par excellence pour non seulement manifester sa place et son rapport à la complétude, mais également altérer, renforcer ou déplacer le rapport de places qui se disposait selon les trois premiers registres. L'interaction verbale forme un tout en soi dans lequel chaque sujet désire prendre part de façon pertinente.

Il va sans dire que le sujet parle à partir de tous les registres en même temps bien qu'il puisse y en avoir un de plus présent.

La parole comme produit idéologique

La valeur discursive d'un mot n'est pas évidente, car chaque sujet s'imagine que la langue est un matériel neutre dont les éléments varient sémantiquement et qu'il peut en disposer à son gré pour organiser et exprimer sa pensée, celle qu'il fait sienne. Admettre, pour un sujet parlant, que

son discours est issu d'une préformation sociale consisterait à nier qu'il est un sujet libre, capable de penser par lui-même, ce qu'il ne peut en toute bonne foi faire. Pourtant, toute parole est produite sur la base d'un système discursif qui vient se coller à la réalité par le biais de représentations sémiologiques. En effet, bien que l'individu identifie naturellement sa langue à la réalité, la réalité qu'il croit percevoir n'est en fait qu'un ensemble de significations, de repères symboliques et de représentations qui forment l'univers sémiologique. C'est dire alors que tout référent est à chercher dans cet univers sémiologique, et non dans la réalité, et que les connotations d'un mot sont le fruit de l'ensemble des traits sémiologiques qu'il recouvre et de leurs relations au système discursif. Le système discursif et l'univers sémiologique sont imbriqués étroitement; le premier ordonne la production du second alors que le second contraint la structure du premier. En fait, le niveau discursif est celui de l'organisation des représentations à partir desquelles le locuteur élabore ce qu'il veut dire et à partir desquelles aussi l'interlocuteur reconstruit ce qu'on a voulu lui dire.

Si la parole est produite par un système discursif, ce dernier est lui-même produit par une idéologie donnée. Toute idéologie se présente sous les formes de la complétude. Effectivement, comme nous l'avons dit plus haut, l'idéologie appréhende, pense et organise la réalité en fournissant un système de termes cohérent et total. Elle ne peut concevoir un extérieur, donc ne peut le reconnaître, à moins que cet extérieur soit un autre système de pensée, alors elle le traduit dans son propre système et lui assigne une place, ce qui, du coup, détruit l'altérité. Le problème avec la complétude, c'est qu'elle n'existe pas. Son paradoxe, "c'est qu'elle se donne pour réalité, bien qu'elle soit cependant toujours prise, et même produite, dans le monde des signes." (Flahault, 1978: 98). En fait,

> "l'interprétation d'un énoncé par son destinataire exige de celui-ci moins un décodage qu'un "calcul", qui reproduit la relation entretenue par l'énoncé avec un certain nombre de repères sélectionnés dans les représentations que l'interlocuteur partage, ou croit partager, avec le locuteur." (Flahault, 1978: 37)

L'illocutoire

Si le contenu de l'énoncé doit être pertinent par rapport à la situation verbale, l'énonciation, elle, doit s'appuyer sur la croyance que l'on a quelque chose à dire de pertinent dans cette même situation. La pertinence d'un énoncé relève de sa visée à la complétude et pour ce faire, le sujet doit non seulement comprendre ce qui est dit, ce pour quoi, ce compte tenu de quoi ou par rapport à quoi c'est dit, mais également croire qu'il est lui-même en mesure de contribuer au tout. Son énoncé, tout comme son énonciation, porte les marques de son propre rapport à la complétude, marques que F. Flahault appelle **insignes**. De plus, comme aucun sujet ne peut juger de sa propre pertinence, il y a toujours, forcément, derrière toute prise de parole une demande sous-jacente de reconnaissance de ce qui se dit. Cette reconnaissance n'est obtenue que si l'autre juge que le sujet parlant remplit bien la condition à laquelle il prétend satisfaire et qu'il porte bien l'insigne qu'il croit porter. C'est parce que la parole comprend toujours cette demande de reconnaissance qu'on la dit **illocutoire**. La force illocutoire relève de l'énonciation et ne se comprend qu'en fonction du rapport de places qui se crée entre les interlocuteurs et le discours. Fondamentalement, ce qui provoque l'illocutoire se résume aux deux questions suivantes: **Qui je suis pour toi?** et **Qui tu es pour moi?**. L'énonciation du locuteur se veut donc un insigne de la place d'où il croit parler.

La variation de l'emploi de JE et de TU comme insigne

La conclusion de l'analyse d'entrevue se résumait à affirmer que JE+ modélisateur et les expressions avec TU correspondaient à des stratégies de réponse. Or, l'acte de poser une question, comme

nous l'avons décrit précédemment, constitue une mainmise sur la parole, le temps et le corps de l'autre; il est donc fondamentalement illocutoire. Non seulement son efficacité repose-t-elle sur la reconnaissance de l'autre comme sujet parlant mais elle se double également de la demande du questionneur d'être reconnu lui-même comme sujet parlant. De plus, toute question place l'autre dans la position d'avoir à répondre. C'est ainsi qu'elle exerce sur l'interlocuteur une contrainte qui dépasse, tout en s'y appuyant, le **qui je suis pour toi** et qui se formule par **ce que tu peux me répondre**. Cette contrainte s'inscrit directement dans les règles discursives qui organisent les possibilités de parole en ce que seuls certains énoncés peuvent être des réponses. Si la réponse n'est pas adéquate, c'est-à-dire si elle ne tient pas compte du contexte d'énonciation, de la place de l'autre, etc., elle ruine le rapport à la complétude auquel l'autre prétend. Dans notre situation d'entrevue, les rôles de questionneur/euse et d'informateur/trice ont été clairement déterminés. L'acte de poser des questions est donc moins violent qu'il ne le serait dans une simple conversation; de même, les réponses visent toutes la complétude de ce tout qu'est l'entrevue, situation à laquelle chaque sujet est assujetti du fait qu'il en a accepté l'avènement. La haute fréquence du JE+ modélisateur dans le discours des femmes ainsi que l'emploi marqué des expressions avec TU dans celui des hommes doivent donc être analysés en fonction de cette place qui se cherche et de ce quelque chose qui se donne à reconnaître.

La force illocutoire de JE+ modélisateur

Le registre le plus facile à identifier est très certainement celui constitué par la présente instance de discours. Les femmes sont autorisées à prendre la parole à titre d'informatrices participant à une entrevue effectuée dans le cadre d'une étude universitaire sur la langue parlée en Estrie. Nos 12 informatrices ayant participé à cet échange verbal, nous pouvons donc assumer qu'elles ont accepté cette place d'interviewée et qu'elles désiraient établir un rapport à la complétude de cette situation.

Je rappelle que chacun/e vise à participer à l'échange verbal de façon pertinente en produisant les preuves (insignes) de sa propre pertinence (de la place d'où il/elle parle et de ce qu'il/elle dit), preuves qu'il/elle soumet à la reconnaissance de l'autre. Dans mon étude, le discours dominant est celui sur la langue. En employant un grand nombre de JE+ modélisateurs, les informatrices ont traduit qu'elles endossaient seules (la présence de JE en est la preuve) l'opinion ou la perception qu'elles avaient sur l'objet dont il était question. Elles ne répondent pas sur l'objet mais sur une attitude par rapport à l'objet. Elles se situent individuellement et explicitement en dehors de ce qui est dit. Elles posent un regard, elles critiquent, elles émettent une opinion ou encore elles disent la perception qu'elles ont de l'objet, et la force illocutoire du JE+ modélisateur réside dans cette individualisation signifiée. Ainsi, la demande de reconnaissance qui sous-tend l'énonciation des femmes est celle du droit d'énoncer, un peu comme si elles demandaient qu'on reconnaisse leur pertinence en tant que sujet plus que la pertinence de leur propos. Ceci rejoint ce que M. de Fornel et P. Encrevé disaient des attitudes propositionnelles qui permettent une reprise, un retour sur ce qui a été dit.

La force illocutoire du JE+ modélisateur réside donc dans le fait que les informatrices, en l'employant, traduisent la conscience qu'elles ont que leur place doit être encore et toujours objet de reconnaissance, et ce avant même la pertinence du propos. C'est dans le registre de la circulation d'insignes que cette force illocutoire se manifeste le plus évidemment. Mais, comme il en a été fait mention, ce registre englobe les trois premiers et, par conséquent, il porte en lui les raisons de cet état de choses, de ce comportement particulier de nos informatrices. En ce qui concerne le premier registre, l'inconscient, il nous est à peu près impossible d'énoncer quoi que ce soit qui s'y rattacherait sans tomber dans une subjectivité qui ici ne serait pas de bon aloi. Au deuxième registre, celui qui s'appuie sur la place du sujet dans la formation sociale à laquelle il appartient, la précarité de cette place provient d'une distribution de la parole basée sur une division sexuelle. De tout temps la parole des femmes, et uniquement celle-ci, a été taxée de bavardage, au point tel qu'une étude menée par V. Aebischer auprès de 60 femmes (45 000 minutes d'interview) a démontré que toutes les femmes reconnaissent le bavardage comme trait racial. Ainsi,

> "c'est une réalité qui associe le bavardage aux femmes. Peu importe que celui-ci soit doté de connotations positives ou négatives, le fait de l'associer à une entité et à tous les membres de cette entité, d'une part, et l'aspect déterministe de cette association, d'autre part, en font un trait racial. Aussi nuancée qu'elle puisse être, la représentation de la femme et du bavardage est de nature raciste et tout le monde, en l'occurrence les femmes, y participe." (Aebischer, 1985: 166)

Certaines endossent ce trait racial, d'autres le rejettent carrément... sur le dos d'autres femmes! Une autre étude de V. Aebischer est basée sur une série de dessins naïfs représentant six femmes et six hommes qui ne se distinguent que par la posture, car il n'y a pas de parole ni d'expression faciale. Ces dessins ont été soumis à une centaine d'étudiant/e/s, 50 femmes et 50 hommes inscrit/e/s en psychologie à l'Université de Paris X, pour que ces dernier/ière/s les décrivent en fonction de ce qu'ils dégagent par rapport à la parole. Les résultats de l'analyse des réponses obtenues sont renversants: **CROIRE, C'EST VOIR**. En effet, étudiantes comme étudiants ont toutes et tous perçu les femmes dessinées comme bavardes, futiles, frivoles, indiscrètes, expressives, communicatives et chaleureuses, et les hommes comme froids, peu communicatifs, peu expressifs, stériles, discrets, sérieux, importants et peu bavards. Pourtant, les 12 dessins représentaient six femmes et six hommes dans des postures identiques. La représentation de la femme bavarde est indépendante des femmes et de leurs parlers. Dans l'étude basée sur les dessins naïfs, les femmes et les hommes ne parlaient pas!

Derrière la capacité de parler intelligemment se trouve la capacité de penser. En qualifiant la parole des femmes de bavardage, on sous-entend qu'elles sont incapables de réflexion et, par conséquent, qu'elles ne peuvent que dire des insignifiances, tenir des propos sans fondement, énoncer des lieux communs... Comment les femmes peuvent-elles être à l'aise dans une interaction verbale comme celle de l'entrevue, où les réponses qu'elles fournissent font l'objet d'une évaluation continuelle, alors qu'elles sont conscientes que la valeur de leur parole est jugée si petite?

Je propose d'interpréter l'usage distinctif du JE+ modélisateur dans le discours de nos informatrices comme la preuve que ces dernières savaient, consciemment ou inconsciemment, que leur parole n'avait pas de poids reconnu et qu'il leur était nécessaire de soumettre leur qualité de sujet à la reconnaissance de l'autre, avant même celle de leur propos. L'emploi du JE marque l'individualité, la non-appartenance à un groupe quelconque. En fait, le seul groupe que connaissent les femmes est celui de l'ensemble de la gent féminine, et comme le bavardage est la caractéristique de la parole des femmes, il était essentiel pour chaque informatrice de se démarquer des autres membres de son groupe si elle désirait obtenir une certaine crédibilité. La forte utilisation des verbes modélisateurs provient du désir de fournir la preuve de la réflexion (de la capacité de penser, de percevoir, d'évaluer, de justifier...) face à l'objet dont il est question. Ainsi, si les femmes de notre corpus ont fait un usage si grand des JE+ modélisateurs, c'est dans un désir de fournir les insignes pertinents de la compréhension qu'elles avaient, d'une part, du système discursif qui leur refuse l'accès à une parole intelligente, le discours des hommes, et, d'autre part, du fait qu'elles étaient interrogées formellement en entrevue, cette entrevue constituant un espace de réalisation qui légitimait leur parole.

La force illocutoire des expressions avec TU

Les hommes de mon corpus ayant accepté de participer aux Recherches sociolinguistiques dans la région de Sherbrooke, ils prennent la parole en tant qu'informateurs. Tous ont accepté de viser la complétude de l'instance de discours de l'entrevue.

Il a été mis en lumière précédemment que les expressions figées SI TU VEUX et TU VAS PRENDRE s'inscrivent dans un schéma d'hypothèse dont l'effet est de solliciter indirectement l'interlocuteur, à titre de n'importe qui, lui ou un autre. Il n'y a aucune demande de sanction du propos

inhérente à l'emploi de telles expressions. Dans le cas de T'SÉ, il est apparu que cette expression équivalait à un **n'est-ce-pas** et qu'elle revêtait trois valeurs différentes selon le contexte dans lequel elle circule: demande de sanction, appuyeur de propos et embrayeur d'opinion personnelle. Tout comme les deux premières expressions, celle-ci comporte une certaine forme de sollicitation indirecte.

Par l'utilisation des expressions idiomatiques, l'informateur place l'enquêteur/trice en situation de partage d'une interaction verbale. La force de SI TU VEUX et TU VAS PRENDRE réside dans le fait que ces dernières font appel à l'enquêteur/trice dans la mesure où il/elle est celui/celle qui pose les questions. Il n'y a aucune demande de reconnaissance ou d'acquiescement au propos tenu, car à ce moment les expressions seraient suivies d'une interrogation et d'une sanction, d'un accord, ce qui n'est pas le cas. Au contraire, d'une certaine façon, l'informateur lui signifie qu'en tant qu'informateur, il cherche à donner des réponses, les meilleures qui soient dans ce contexte-ci, mais ce n'est pas parce qu'il ne trouve pas les mots exacts, ou que sa réponse n'est pas totale, qu'il n'est pas pertinent comme locuteur. Il joue à l'entrevue mais sa parole ne saurait se réduire à celle de l'entrevue et la seule reconnaissance qu'elle peut lui accorder est celle de sa liberté de sujet parlant. C'est comme s'il y avait une inversion des demandes de reconnaissance. L'informateur rappelle constamment à son interlocuteur/trice son rôle enquêteur, ou plus exactement, qu'il/elle est momentanément, par un concours de circonstances, son interlocuteur/trice, et à quelque part celui/celle qui a un certain pouvoir sur ce qu'il dit, mais, par la valeur non-définie du TU des expressions, il lui signifie aussi qu'il/elle ou un/e autre serait à même de recevoir et de concevoir son propos.

La force illocutoire de T'SÉ se manifeste dans le registre du discours dominant mais provient en réalité du registre idéologique. Qui ne connaît pas l'expression **T'SÉ, veux dire**? Un peu comme si on disait la non nécessité de détailler plus en profondeur du fait que de toute façon on se comprend. Donc, encore une fois, l'emploi du T'SÉ chez les hommes témoigne de leur croyance à un univers commun et de la justesse de la place qu'ils y occupent. La force illocutoire de T'SÉ réside également dans la place qu'il assigne à l'interlocuteur, une place de témoin du caractère personnel de ce qui va être dit, personnel mais non exclusif. En fait, si l'autre ne reconnaît pas que la parole est un fait d'hommes, c'est que lui-même n'est pas à sa place, qu'il signifie sa non-appartenance à un univers sémiologique supposé commun dans lequel les hommes ont toujours manié la parole en maîtres, et du même coup, il ne peut viser la complétude de l'idéologie d'origine.

Que la parole soit un acquis pour les hommes me semble le noeud central du système discursif d'où parlent nos informateurs, et les hommes en général. Bien sûr, il existe des hommes pour qui le maniement de la parole est plutôt difficile, mais nous croyons qu'il faut y voir la crainte de ne pas être à la hauteur des exigences de l'acquis plus que le sentiment d'insécurité face à sa propre pertinence.

Je propose d'interpréter la haute fréquence des expressions avec TU dans le discours de mes informateurs comme la manifestation de leur désir de renvoyer à l'autre que, peu importe d'où ils parlent, la parole leur est acquise. Les TU VAS PRENDRE, SI TU VEUX, et particulièrement le T'SÉ, sont une projection sur l'autre des croyances que les hommes entretiennent par rapport à leur droit de parler. Non seulement ne se posent-ils pas comme sujet explicite de leur dire, et s'ils le faisaient, ils reconnaîtraient que l'acquis n'en est pas un, mais en plus ils s'octroient le droit de faire comprendre à l'interlocuteur qu'ils sont les auteurs de la place qu'il occupe. En effet, la valeur fondamentale du TU revient à attribuer une place particulière à l'autre. Bien que le TU des expressions ait une valeur non-définie, il possède encore la marque d'une allocution, aussi indirecte qu'elle puisse être et qui se traduit par **toi ou un autre**. Le discours des hommes de mon corpus en est un de dominant. En effet, "c'est bien le propre du **discours du pouvoir** que celui qui le tient occupe la place de l'instance qui reconnaît et assigne les places." (Flahault, 1978: 210) S'il en est ainsi, c'est que la parole a toujours été aux mains des hommes. En fait, les hommes sont responsables de la majorité des systèmes de pensée à l'intérieur desquels ils règnent en maîtres absolus de la parole.

Conclusion

L'approfondissement de la question de la parole a consolidé ma conviction que les unités linguistiques n'apparaissent pas n'importe comment dans n'importe quel contexte avec n'importe qui indifféremment du sexe des interlocuteur/trice/s en présence, et que faire abstraction de l'extra-linguistique ne peut que conduire à une impasse dans la compréhension de la parole. L'explication que je cherchais se trouve dans la perception du deuxième registre identifié par F. Flahault, celui basé sur les systèmes discursifs qui délimitent la place du sujet selon le sexe et la formation sociale à laquelle il appartient. Mes interprétations, qui sont elles-mêmes inscrites dans un système discursif produit par une idéologie, sont à la hauteur de ma propre perception de ce qui se véhicule dans le deuxième registre, c'est-à-dire des idéologies qui se vivent dans notre société, principalement celles qui touchent, de près ou de loin, à la place des femmes et des hommes dans la parole. En arriver à un constat de rapport de force entre les deux sexes n'est pas étonnant, car non seulement le rapport de force constitue-t-il cette violence du réel qu'il nous faut tempérer et régulariser, mais il renvoie également à une idéologie qui lui assure une survie dans nos systèmes de pensée et qui attribue les places de façon à rendre la vie tout de même tolérable... à condition que chacun et chacune respecte l'attribution des places. Enfin, aucune interprétation n'est totale, elle est plus ou moins plausible en fonction de l'ampleur de la recherche dans divers domaines en vue de cerner le pourquoi d'un phénomène en particulier, dans mon cas, le comportement différentiel des femmes et des hommes en regard de l'emploi des pronoms personnels JE et TU.

BIBLIOGRAPHIE

AEBISCHER, Verena (1985), *Les femmes et le langage*, Coll. "Sociologie d'aujourd'hui", P.U.F., Paris.

BEAUCHEMIN, Normand (1972), *Les enquêtes sociolinguistiques dans la région de Sherbrooke*, Disponibles et accessibles sur le système informatisé MUSIC de l'Université de Sherbrooke, Sherbrooke.

BEAUCHEMIN, Normand (1983), *Quelques différences entre le vocabulaire des hommes et celui des femmes en québécois parlé*, Miméo, Université de Sherbrooke, Sherbrooke.

DE FORNEL, Michel et ENCREVE, Pierre (1983), "Le sens en pratique: construction de la référence et structure sociale de l'interaction dans le couple question/réponse", dans *Actes de la recherche en sciences sociales*, n° 46, mars 1983, pp. 1-28.

DESHAIES, Denise (1982), *Référence personnelle et type de discours en situation d'entrevue*, Miméo, Université Laval, Québec.

FLAHAULT, François (1978), *La parole intermédiaire*, Seuil, Paris.

LABERGE, Suzanne (1977), *Étude de la variation des pronoms personnels définis et indéfinis dans le français parlé à Montréal*, Thèse de doctorat non-publiée, Université de Montréal, Montréal.

SIMONIN, Jacky (1977), "Analyse interactionnelle du système question/réponse dans la pratique d'interview", dans *Travaux de recherches sémiologiques*, n° 25, janvier, pp. 13-47.

L'EMPLOI DES MODES INDICATIF ET SUBJONCTIF DANS LE FRANÇAIS PARLÉ DE LA VILLE DE QUÉBEC[1]

Julie Auger
Université Laval

Cet article présente les résultats d'un mémoire de maîtrise intitulé "L'emploi des modes indicatif et subjonctif dans les propositions subordonnées dépendant d'une structure impersonnelle dans le français parlé de la ville de Québec". Je tâcherai de faire ressortir les résultats les plus intéressants, tant en ce qui concerne la description de l'usage qu'en ce qui a trait aux facteurs linguistiques et sociaux qui influencent cet usage.

Pour réaliser cette étude de l'emploi des modes dans les propositions subordonnées dépendant d'une structure impersonnelle, j'ai choisi de constituer mon corpus à même les données de l'enquête sur le "Français parlé dans la ville de Québec" réalisée sous la direction de Mme Denise Deshaies. Comme cette enquête implique environ 250 heures de langage oral, j'ai préféré travailler sur un sous-échantillon servant déjà de base à diverses études sur les pronoms personnels. Ce sous-échantillon comprend les entrevues individuelles de vingt adolescentes et adolescents des quartiers St-Sauveur et Ste-Foy, de même que les entrevues réalisées avec leurs parents, ce qui totalise 53 informateurs.

Le but de cette étude était double: il s'agissait d'une part de décrire l'emploi des modes avec les différentes structures impersonnelles et, d'autre part, d'identifier les facteurs linguistiques et sociaux qui sont à la base de la variation modale. Pour la plupart des structures, l'étude était surtout descriptive puisque ces structures entraînent l'emploi exclusif d'un seul mode. Dans certains cas, cette absence de variation pouvait être due au petit nombre d'occurrences fournies par le corpus; c'est le cas, par exemple, de structures telles que IL S'AGIT ou C'EST DÉCIDÉ, structures pour lesquelles je n'ai recueilli qu'un exemple. Dans d'autres cas, sans toutefois pouvoir affirmer que la variation modale est impossible avec telle ou telle structure, les exemples recueillis indiquent que, dans le style de discours étudié, l'emploi de tel ou tel mode est nettement dominant. Une telle situation a été observée avec la structure C'EST RARE: les 59 occurrences relevées contiennent toutes des verbes subordonnés à l'indicatif ou des verbes non-marqués, c'est-à-dire ces formes verbales pour lesquelles il n'est pas possible de savoir s'il s'agit d'un indicatif présent ou d'un subjonctif.

(1) "Tous les soirs après souper, euh... c... c'est rare qu'on sortait pas." (105: 48)[2]

(2) "Si une famille fait du sport, c'est ben rare qu'i (il)[3] va n' (en) avoir qui vont être à côté d'la coche." (205: 188)

---

[1]Cette communication est le résumé d'un mémoire de maîtrise réalisé avec l'aide d'une bourse spéciale de maîtrise décernée par le Conseil de recherches en sciences humaines du Canada (C.R.S.H.C.) pour l'année 1986-87.

[2]Le premier nombre correspond au numéro du locuteur et le deuxième correspond à la page de la transcription.

[3]Ce qui est noté entre parenthèses n'a pas été prononcé par l'informateur; il s'agit soit d'un équivalent en français standard d'un élément prononcé, soit d'un élément qui a été élidé.

J'ai cependant relevé un exemple produit par un étudiant lors d'une communication; il s'agit donc d'une situation très formelle, style de discours qui n'est pas représenté dans notre corpus, et cet exemple contient un subjonctif.

(3) "C'est assez rare qu'on fasse ça" (Homme, 30 ans, Québec)

Le cas de C'EST RARE est particulièrement intéressant du fait qu'il s'agit de l'une des rares structures (sinon la seule dans notre corpus) pour lesquelles le nombre d'exemples permet d'affirmer que le français québécois parlé en style informel se distingue radicalement de la norme décrite par les grammairiens: ceux-ci, de même que les auteurs qui ont étudié l'emploi des modes dans la langue écrite, ont en effet observé l'emploi exclusif du subjonctif. Notons toutefois que l'emploi de l'indicatif est aussi connu en français parlé de l'Hexagone puisque sur onze occurrences, Nerio (1978) a recueilli 3 indicatifs.

Ces phénomènes catégoriques ne sont généralement pas envisagés dans les études variationnistes mais j'ai choisi d'en tenir compte dans mon mémoire principalement en raison du fait qu'il n'existe aucune étude, descriptive ou variationniste, de l'emploi des modes en français québécois. Cette prise de position m'a aussi amenée à considérer le comportement modal dans les subordonnées dépendant de structures impersonnelles propres au français québécois telles que ÇA ADONNE et C'EST PLATE.

(4) "ça adonnait qu'on pouvait partir toutes les deux pis euh... y aller, tsé." (105: 17)

(5) "Ben j'veux dire que c'est... c'est plate que l'gars i (il)... tsé i (il) fait tout brusque" (045: 182)

Dans ces cas, il faut cependant admettre qu'en raison du petit nombre d'occurrences, je ne peux prétendre donner une description complète de ces structures. Ce type d'étude a le mérite de constituer une première description, un travail de débroussaillement et il permet de se poser un certain nombre de questions concernant l'emploi des modes avec ces structures. Les exemples de ÇA ADONNE permettent par exemple de se demander si le sens de cette structure n'est pas lié au mode employé dans la subordonnée: j'ai en effet relevé des cas où l'emploi de l'indicatif confère à ÇA ADONNE un sens proche de celui de ÇA ARRIVE, structure qui se fait suivre surtout de l'indicatif dans notre corpus alors que dans d'autres cas, où le verbe subordonné est malheureusement non-marqué, le sens de cette expression se rapproche davantage de "l'occasion se présente". Il faudrait donc poursuivre les dépouillements et tenter de vérifier cette hypothèse. Je pense qu'en ce moment, une étude purement variationniste de certains phénomènes précis de l'emploi modal n'aurait donné qu'une vue partielle du problème étudié.

Dans le corpus, six structures ont entraîné de la variation modale en subordonnée. Dans le cas de C'EST UNE BONNE AFFAIRE / C'EST UNE BONNE CHOSE, de ÇA SE PEUT et de ÇA ME SURPREND, le subjonctif est le mode le plus employé mais dans chaque cas, j'ai recueilli une occurrence d'indicatif. Le nombre d'occurrences de ces structures était insuffisant pour permettre une analyse quantitative à l'aide de tests statistiques ou d'analyses probabilistes mais ces trois occurrences d'indicatif donnent déjà une idée des facteurs linguistiques et sociaux qui ont une incidence sur le choix du mode. Voici ces trois exemples de même que trois exemples au subjonctif pour comparer:

(6) "(il) y a des, (il) y a des... comme le, l'histoire sainte pis catéchisme ça c'est une bonne affaire qu'y (ils) ayent (aient) débarqué ça." (214: 314-15)

(7) "Mais... euh... c'est une bonne affaire qu'une personne... va à (l') école pis qu'y (il) étudie bien, pis qu'y (il) apprend" (141: 93)

(8) "après la mort ça s'peut pas que l'monde aille en enfer" (144: 331)

(9) "Des fois j'y allais pas l'matin, j'étais là dans l'après-midi, ça s'pouvait qu'i (il), qu'i (il) m'voyait l'après-midi" (036: 43)

(10) "D'abord, (elle) est assez renseignée que j'pense pas qu'a (elle) soit enceinte, ça m'surprendrait qu'a (elle) soit enceinte." (148: 240)

(11) "Ah m... ça serait, ça s... ça m'surprendrait qu'i (il) ferait un enfant lui, ah!" (104: 314)

Dans le cas de la structure C'EST UNE BONNE AFFAIRE, aucun facteur linguistique ne semble expliquer la variation modale observée mais si l'on se penche sur les facteurs sociaux, on constate que l'informatrice qui a produit cet exemple est l'une des moins instruites du sous-échantillon: elle n'a fréquenté l'école que pendant deux ans. Cette informatrice manifeste de surcroît une tendance à employer l'indicatif dans la subordonnée dépendant de C'EST NORMAL. Je reviendrai sur cette structure un peu plus loin.

Dans le cas des structures ÇA SE PEUT et ÇA ME SURPREND, il semble que ce soient d'abord des facteurs linguistiques qui influencent la variation: dans les deux cas, en effet, on constate une concordance entre les verbes de la principale et de la subordonnée, soit deux imparfaits avec ÇA SE PEUT et deux conditionnels avec ÇA ME SURPRENDRAIT. Une telle tendance à employer deux conditionnels n'est pas particulière au français québécois puisqu'elle a été observée par plusieurs auteurs qui ont étudié le français européen. Grevisse (cité dans Cohen, 1965), Tanase (1943), Molen (1923), Damourette et Pichon (1911-40) et Cohen (1965) ont observé ce phénomène de concordance mais ils l'attribuent soit au "menu peuple", soit à des parlers régionaux. En ce qui a trait à l'emploi de l'imparfait dans des subordonnées où l'on s'attendrait à voir apparaître le subjonctif, Togeby a observé que dans certains contextes, le locuteur doit recourir à l'imparfait de l'indicatif pour indiquer que l'événement subordonné appartient à l'époque passée et c'est ce que l'observe dans des phrases comme "Il semble que Voltaire comptait sur ces hésitations" (Togeby, 1974: 11) ou, comme dans l'exemple (1), "Tous les soirs après souper, euh... c... *c'est rare qu'on sortait* pas." (105: 48). Il a également noté des cas où l'on emploie l'imparfait en subordonnée par concordance avec le verbe de la principale dans des exemples tels que "Il semblait qu'on pouvait entendre les bruits minuscules" (Togeby, 1974: 13). Nerio (1979) a relevé de tels exemples dans le français parlé de l'Hexagone mais il attribue de tels cas de concordance à des locuteurs qui n'appartiennent pas aux classes dominantes de la société.

Cet emploi de l'imparfait et du conditionnel en subordonnée, là où la norme recommande l'emploi du subjonctif, n'affecte pas que ces deux structures. J'ai relevé de tels cas de concordance de deux imparfaits avec C'EST RARE (dans ce cas, l'imparfait est parfois employé avec un présent en principale et il est également employé par concordance avec un autre imparfait), C'EST FRÉQUENT, ÇA ARRIVE, ÇA ADONNE, ÇA VIENT PAS À L'IDÉE. En ce qui a trait à la concordance au conditionnel, ce facteur explique la presque totalité des subordonnées à l'indicatif avec la structure IL FAUT.

Dans le cas de IL FAUT, j'ai recueilli 694 occurrences et il a donc été possible de soumettre ces résultats à l'analyse probabiliste de Varbrul. Cette analyse révèle que le facteur dont l'influence est la plus décisive est le temps de la proposition principale: le conditionnel favorise fortement l'emploi de l'indicatif alors que le présent et l'imparfait favorisent l'emploi du subjonctif. Du côté des facteurs sociaux, seul le sexe a été retenu: les femmes favorisent l'emploi de l'indicatif alors que les

hommes prévilégient le subjonctif. L'analyse des cas où IL FAUT est au conditionnel démontre de plus que ce sont les gens les moins instruits et ceux qui habitent le quartier St-Sauveur, quartier moins favorisé, qui emploient le plus le conditionnel dans ce contexte.

(12) "Cette année, (il) faudrait (que) j'm'arrangerais, j'aimerais ça être, être brigadière. (Il) faudrait (que) j'appellerais Aimé St-Laurent" (104: 343)

(13) "Ah non! moi si i (il) faudrait que j'me mettrais à parler euh..." (110: 324)

Cet emploi du conditionnel représente 40% des cas où IL FAUT est au conditionnel mais, sur l'ensemble des données du corpus, cet emploi représente 5,14% des occurrences. L'emploi du subjonctif dans ce contexte reste donc très majoritaire: 93,83%, si l'on ne tient pas compte des verbes non-marqués.

(14) "(il) va falloir qu'i (il) soye (soit) capable d'es (les) comprendre pis d'leur parler" (245: 545)

(15) "(il) Faut surtout qu'i (il) soye (soit) beau pis qu'i (il) soye (soit) fin." (001: 25)

Le corpus a également fourni 5 occurrences de présent ou d'imparfait de l'indicatif mais ces cas ne représentent que 1,28% des 694 occurrences du corpus, ce qui semble donc assez exceptionnel en français québécois.

(16) "Ben supposons e... tu fais une lettre à (la) dactylo, ben, (il) faut qu'ça sert comment parler, comment lire, écrire... Ça sert" (028: 154)

Les résultats concernant l'incidence des facteurs sociaux sur le choix du mode avec IL FAUT s'opposent à ce que l'on observe avec ÇA ARRIVE: dans ce dernier cas, ce sont les gens les moins instruits qui favorisent l'emploi du subjonctif. J'ai recueilli 59 occurrences de cette structure et celles-ci indiquent que l'indicatif est le mode le plus employé dans ce contexte: 45 indicatifs, 6 subjonctifs et 8 verbes non-marqués. C'est dans le cas de cette structure que l'on peut observer l'incidence du plus grand nombre de facteurs linguistiques: Varbrul indique en effet que le facteur le plus nettement subjonctivisant est la présence du verbe POUVOIR dans la principale.

(17) "tu acceptes pas de verre d'un autre, ça peut arriver qu'(il) y aye (ait) d'la drogue dedans" (456: 132-33)

(18) "ça peut arriver moé que j'dise quelque chose de travers" (145: 662)

Il ressort ensuite que le passé composé en principale est le seul temps qui favorise l'emploi du subjonctif en subordonnée; le présent favorise pour sa part l'emploi de l'indicatif alors que l'imparfait, le futur périphrastique et le conditionnel passé entraînent l'emploi catégorique de l'indicatif.

(19) "si c'est par rapport à moé, j'vas parler, tsé comme ç'a déjà arrivé (que) ça soye (soit) par rapport à moé j'm'en suis mêlée" (045: 13)

(20) "Pour dire qu'on s'est tenu ben, ben en groupe, là, euh... ça arrivait, des fois, qu'on jasait au coin, au restaurant" (141: 54)

(21) "Une supposition qu'a (elle) dit un mot (...) ça va arriver qu'on va y (lui) dire mais pas pour dire insister tellement"

Pour cette structure, nous avons également tenté de déterminer l'incidence de la présence d'un pronom objet dans la principale: le nombre d'exemples est insuffisant pour conclure mais il est intéressant de constater que les pronoms ME et TE n'entraînent que l'indicatif en subordonnée alors que le caractère formel du vouvoiement dans les questions des enquêtrices aux parents de l'enquête semble favoriser l'emploi du subjonctif.

(22) "Même moi ça m'arrive des fois que j'vas être fâchée là, puis euh... j'vais sacrer" (156: 178)

(23) "Vous, est-ce que... ça vous arrive quand Michel, là, fait quelque chose de pas correct, que vous euh... lui disiez, que vous..." (MG dans corpus 156: 166)

J'ai parlé plus tôt de l'emploi de l'imparfait et du conditionnel dans des subordonnées où ces temps entrent en concurrence avec le subjonctif. L'étude de ÇA ARRIVE m'a amenée à considérer que le futur périphrastique est un autre temps de l'indicatif qui manifeste une certaine tendance à entrer aussi en variation avec le subjonctif. Les exemples (19) et (20) présentent de tels cas d'emploi de ce temps.

Cet emploi du futur périphrastique m'amène enfin à parler de la structure C'EST NORMAL. Contrairement à ce que tous les auteurs ont observé, cette structure entraîne une importante variation modale en français québécois puisque j'ai recueilli presqu'autant d'indicatifs que de subjonctifs: 9 indicatifs et 11 subjonctifs. Six des 9 indicatifs sont des futurs, futur périphrastique à la forme affirmative et futur simple à la forme négative, et 5 d'entre eux ont été produits par un locuteur qui emploie ce temps de façon catégorique dans ce contexte. Les autres indicatifs sont des présents.

(24) "lui, i (il) va à (l') école, c'est normal (que) ça va arriver" (026: 59)

(25) "j'vas faire un dessin pis c'est normal qu'i sera pas parfait" (026: 69)

(26) "c'est normal que c'est pas toute, toute la même chose, on peut pas être tous euh... pareil non plus" (141: 281)

Dans le premier exemple, la conjonction QUE est élidée et je me suis alors basée sur l'intonation pour décider qu'il s'agit d'une subordonnée et non d'une indépendante juxtaposée. Cette élision de la conjonction QUE est un phénomène relativement répandu en français québécois parlé, comme le révèlent les études de Bickerton (1972), Cedergren et D. Sankoff (1974), G. Sankoff (1973 et 1974).

Plusieurs locuteurs emploient les deux modes dans la subordonnée dépendant de C'EST NORMAL, certains emploient exclusivement le subjonctif alors qu'un seul emploie l'indicatif de façon catégorique. Aucun facteur social ne semble caractériser cette variation modale.

Voici donc, de façon très rapide, les principaux résultats obtenus dans cette étude. Il apparaît que, pour les structures que j'ai pu étudier, l'incidence des facteurs linguistiques est la plus impor-

tante pour le choix du mode en subordonnée. Bien que nous ne l'ayons pas mentionné explicitement dans cet exposé, c'est d'abord la valeur lexicale des structures impersonnelles qui exerce son influence sur le choix du mode. L'imparfait, le conditionnel et le futur (surtout périphrastique) sont les temps les plus employés en concurrence avec le subjonctif. La présence du verbe POUVOIR avec la structure ÇA ARRIVE est le facteur le plus déterminant dans ce cas. Quant aux facteurs sociaux, j'ai pu observer leur influence avec les ÇA ARRIVE et IL FAUT mais alors que pour ÇA ARRIVE, ce sont les gens les moins instruits qui favorisent l'emploi du subjonctif, j'ai constaté la tendance contraire avec IL FAUDRAIT. Il semble donc pour l'instant qu'on ne puisse pas généraliser l'influence des facteurs sociaux à tous les contextes d'emploi du subjonctif.

Je n'ai malheureusement pas pu faire le tour de tous les résultats obtenus lors de cette recherche mais j'ai essayé de faire ressortir les plus importants.

BIBLIOGRAPHIE

BICKERTON, Dereck (1972), "Quantitative versus dynamic paradigms: The case of Montreal QUE", dans C.-J. N. Bailey et R.W. Shuy (éds), *New ways of analyzing variation in English*, Georgetown University Press, Washington, D.C., pp. 23-43.

CEDERGREN, Henrietta et SANKOFF, David (1974), "Variable rules: Performance as a statistical reflection of competence", dans *Language*, 50, pp. 333-355.

COHEN, Marcel (1965), *Le subjonctif en français contemporain*, Centre de documentation universitaire, Paris.

DAMOURETTE, J. et PICHON, E. (1911-1940), *Des mots à la pensée. Essai de grammaire de la langue française*, Éditions d'Artrey, Paris.

MOLEN, Willem van der (1923), *Le subjonctif: sa valeur psychologique et son emploi dans la langue parlée*, Druk van de N.V., Van de Garde et Co's, Drukkerij, Zalt-Bommel.

NERIO, Tsuysoshi (1978), *Problèmes et méthodes de la syntaxe du subjonctif en français contemporain: Ébauche pour une étude sociolinguistique*, Thèse de doctorat de 3^{e} cycle, Université de Tours.

SANKOFF, Gillian (1973), "Above and beyond phonology in variable rules", dans C.-J. N. Bailey et R.W. Shuy (éds), *New ways of analyzing variation in English*, Georgetown University Press, Washington, D.C., pp. 44-66.

SANKOFF, Gillian (1974), "A quantitative paradigm for the study of communicative competence", dans R. Bauman et J. Sherzer (éds), *Explorations in the ethnography of speaking*, Cambridge University Press, pp. 19-49.

TANASE, Eugène (1943), *Essai sur la valeur et les emplois du subjonctif en français*, Imprimerie A. et F. Rouvière, Montpellier.

TOGEBY, Knud (1974), "Emploi du temps aux dépens du mode", dans *Studia Neophilologica*, Upsal, t. 46, n^{o} 1, pp. 10-14.

A STUDY OF THE (-ING) VARIABLE IN THE OTTAWA MALE COMMUNITY

Dianne Fai
Université d'Ottawa

This pilot study was undertaken last semester as a class project to demonstrate Variation Theory. Preliminary work was initiated in an area of my interest, which is to determine if actual differences in linguistic behaviour exist across the continuum of the male social group, specifically between heterosexual and homosexual males; as well as to observe their speech patterns and decide whether further research is justified.

Intuitively, differences do exist. In fact, they have reached the stage of social comment as evidenced by the prevalence of stereotyped attitudes. This project was based on the available literature concerning female linguistic behaviour.

The assumption is that homosexuals have very different social networks from their heterosexual conterparts. If these are less dense and multiplex, and if homosexuals function as "lames" within the speech community (i.e. as defined by Labov (1972), those outside of the central group and its culture), they, then, should display linguistic behaviour similar to that of women. Quantitative sociolinguistic analysis, therefore, was indicated to investigate the true speech patterns within these two special groups.

An interview schedule was tailored pertaining specifically to the male communities that were to be studied. The schedule was designed to elicit a large body of vernacular speech dealing with aspects of sexuality, lifestyles, moral issues, religion, and death. One homosexual and one heterosexual male were chosen to conduct two interviews each, using informants from their own social networks.

Although both interviewers were instructed to utilize the interview schedule only as a guide, and to participate equally in the conversation, the heterosexual interviewer used the schedule exclusively, and his collected interviews revealed a more formal speech style.

The homosexual interviewer, on the other hand, was very adept, eliciting three hours of vernacular speech, with very little attention being paid to the recording equipment. Unfortunately, he was only able to collect one interview. During their discourse, the two homosexuals appeared to unconsciously switch roles. For this reason, the interviewers were included in the sample population, thereby, increasing it by two.

In this study, the selected variable was the suffix (-ing) which is used to form participle endings, and verbal nouns. As well, it appears in certain pronouns, such as "something, anything", et cetera; and it has the realized values of [in] and [iŋ] in English. What is crucial is the presence or absence of velarity in the final nasal.

This variable was chosen as the logical place to begin the search for differences as it is both well-documented and historically attested. The assumption was that homosexuals would display a more conservative speech style. That is, they would use more of the variable (-ing). Interestingly enough, the results rejected the hypothesis.

The dropping of the "-g" has covert prestige within the male social network, as noted by Trudgill (1974). Group solidarity exerts a normative pressure towards maintaining its own vernacular culture.

According to Milroy (1980), individual linguistic variability strongly correlates with one's degree of integration into local social networks and the community. Based on the results of this small sample, it appears that it is the homosexuals who have the stronger social networks and display covert prestige, as they utilized the variant [in] more often.

TABLE 1

Indices of [in] usage

	Fischer	Labov	Fai*
Least formal	63	50	88%
Intermediate	52	31	64%
Most formal 03	13	44%	

* Subjective judgement by analyst.

NB: A high index score indicates more usage of the variant [in]. This applies to all of the tables found in this paper.

The index scores were arrived at by calculating the straight percentage of the number of times that words displayed the value [in], divided by the total number of words where either value was possible. A high index score reveals greater usage of the non-standard form, or more of the variant [in]. Index scores were calculated, first for each informant, then, for each word class (i.e. verbs, pronouns, nouns, and adjectives), and finally, for each of the following phonological environments (i.e. a voiced consonant, a voiceless consonant, a vowel, and a pause). Index scores were, then, subsequently calculated for each social group (i.e. heterosexual and homosexual), and finally, for the two conversational styles elicited (i.e. casual and careful).

According to Wald and Shopen (1983: 519), verbs tend to allow the [in] variant more (see Table 2). The next most likely hosts are pronouns, followed by nouns and adjectives that are usually derived from the verbs. However, note that this study of male speech did not conform to the general pattern as evidenced in the literature. The verbs revealed the greatest tolerance of [in], with adjectives as the next most likely hosts, followed by nouns. The pronouns revealed the least amount of usage of the variant [in]. These results are not particularly revealing, however, due to the small numbers of nouns and adjectives that appeared in the corpus. Also, the heterosexual who demonstrated the most formal style of speech used 51 of the 100 pronouns.

Some linguistic environments appear to provide constraints on the rule and can be ranked in the order in which they favour the application of the rule (see Table 3). The most favoured environment for the variant [in] to appear is when the variable is followed by a voiced consonant. The next is the voiceless consonant. The least favoured environments for appearance of [in] are where the variable in question is followed by a vowel or a pause.

TABLE 2

Index scores for each word class

	[in]	Total (-ing)	Index Score
Verbs	280	379	74%
Pronouns	30	100	30%
Nouns	15	31	48%
Adjectives	30	55	55%

TABLE 3

Index scores for each linguistic environment for all informants

	[in]	Total (-ing)	Index Score
_#C[+vce]	82	117	70%
_#C[-vce]	126	201	63%
_#V	124	208	60%
_#	23	39	59%

This study was able to clearly differentiate between the two styles of casual and careful speech, and to support previous research. For casual speech, the total number of (-ing) words was 323, with 283 words demonstrating usage of the value [in], for an index score of 74% (see Table 4). For careful speech, the total number of (-ing) words was 242, with 117 words receiving the [in] pronunciation, for an index score of 48%.

TABLE 4

Total number of (-ing) words

Stratification by style*

	[in]	Total (-ing)	Index Score
Casual	283	323	74%
Careful	117	242	48%

* Subjective judgement by analyst.

The next calculations, shown on Table 5, involved collapsing the speakers into two groups based on sexual preference. For the two homosexuals, the total number of (-ing) words used was 198. Out of these, 152 words received the [in] value, for an index score of 77%. For the three heterosexuals, the total number of (-ing) words used was 367, with 203 receiving the [in] pronunciation, for an index score of 55%.

TABLE 5

Total number of (-ing) words

Stratification by sexual preference

	[in]	Total (-ing)	Index Score
Homosexual	152	198	77%
Heterosexual	203	367	55%

Again, calculations based on this division of sexual preference were done for all of the word classes and for all of the given linguistic environments. On Table 6, the resulting stratification clearly reveals that the homosexual males used more of the variant [in] in all of the words generally, and specifically in the verb and noun classes, but less in the adjective class. Both groups used the same amount of the [in] variant in the pronoun class. On Table 7, stratification by linguistic environments also reveals that the homosexual speakers used more of the [in] variant than the heterosexuals, except when the word's following environment is a pause. Here, it is the heterosexual males who used more of the variant [in], in fact, crossing over the homosexuals' index score.

TABLE 6

Index scores for each word class

	Verbs	Pronouns	Nouns	Adjectives
Homosexual	91%	29%	80%	47%
Heterosexual	64%	30%	33%	65%

TABLE 7

Index scores for each following environment

	_#C[+vce]	_#C[-vce]	_#V	_#
Homosexual	79%	78%	79%	63%
Heterosexual	65%	56%	49%	67%

Although the index scores are interesting, the results are hardly revealing. It is obvious from the stylistic problems that arose in this study that only one interviewer should be utilized in future data collection, in order to control for style. Age and socio-economic factors would also be relevant to this research material.

Regardless, the collected corpus did evince lexical, phonetic, and phonological variation. Differences in intonation curves, vowel length, and stress emphasis were noted in the interviews, hence, this undertaking does deserve further research and analysis. For my experimental phonetics project next semester, I intend to explore and to measure these speech qualities.

I would like to thank Shana Poplack for her encouragement and guidance, the interviewers and their informants, and AEDILL for making this conference possible.

BIBLIOGRAPHY

COATES, Jennifer (1986), *Women, men, and language*, Essex, Longman Group UK Limited.

DOWNES, William (1984), *Language and society*, London, Fontana Paperbacks.

FISCHER, John (1958), "Social influences in the choice of a linguistic variant", in *Word*, 14, pp. 47-56.

LABOV, William (1966), *The social stratification of English in New York City*, Washington, Centre for Applied Linguistics.

LABOV, William (1972), *Sociolinguistic patterns*, Philadelphia, University of Pennsylvania Press.

MILROY, L. (1980), *Language and social networks*, Oxford, Basil Blackwell.

TRUDGILL, Peter (1974), *The social differentiation of English in Norwich*, Cambridge University Press.

WALD, Benji and SHOPEN, Timothy (1985), "A researcher's guide to the sociolinguistic variable (-ing)", in *Language: Introductory readings*, Eds. V. Clarke and R. Eschholz, 4th edition, New York, St. Martin's Press.

AN APPLICATION OF CORRESPONDANCE ANALYSIS TO THE MULTIPLE VARIANT PROBLEM

Wladyslaw Cichoki
University of New Brunswick

L'approche classique à la quantification des variables sociolinguistiques est la méthode d'"index pondéré" (weighted-index method) (Labov 1966). Cette méthode présente pourtant des difficultés pour ce qui est de l'analyse des variables qui ont trois variantes ou plus. Dans cet article nous voulons illustrer une solution à ce problème en termes d'analyse des correspondances (ou "Dual Scaling"), une technique statistique multivariée qui présente sous forme graphique la structure d'une variable. Nous verrons à partir des données de l'anglais parlé à Glasgow et de l'anglais parlé à Toronto quelques avantages de cette approche.

The Problem[1]

The basic approach to the quantification of linguistic variation has been to calculate subject scores using the weighted-index method (Labov 1966). An often cited example is the (eh) vowel in New York English: the five phonetic realisations are ordered along a continuum of phonetic height, which also corresponds to the social continuum of standard/non-standard.

(eh - 1) [I$^{\varepsilon}$:ə]
(eh - 2) [e$^{\varepsilon}$:ə] [ɛ$^{\varepsilon}$:ə]
(eh - 3) [æ$^{\wedge}$:]
(eh - 4) [æ$^{\wedge}$:]
(eh - 5) [a:]

To arrive at a subject's score the index assigned to a variant is multiplied by the proportion of tokens of that variant used by the subject, and these products are added together. The higher the score the closer a subject's pronunciation is to the lower vowel and, in this case, to the standard. The subject scores are used subsequently to establish correlations with linguistic and social constraints.

The method appears to be straightforward, but underlying it are statistical assumptions that need to be stated explicitly. The main assumption is that the variants are considered to be interval-type data: i.e. they can be arranged in a linear order with an equal distance or weight between the successive points of this order. In many cases there is a linguistic or social factor which is inherently gradient and thus provides a basis for linearly ordering the variants. There are cases, however, for which there is no continuum on which to establish an order. Scottish English (r), for

[1]This paper is a revised version of Cichoki (1986, Chapter 4).

example, has three phonetic variants - [r, ɹ] and ∅ - among which there is no phonetic gradient and no unequivocal social gradient (Romaine 1978). In Belfast English (a) there is a curvilinear correspondence between social and phonetic facts: the local prestige form associated with the middle class is [a] but working class speakers use [ɛ] as well as [a, ă,ɔ] (Milroy 1982). Headless relatives in Montréal French have three variants - *ce que*, *qu'est-ce que* and *qu'osque* - which cannot be ordered (Kemp 1979). Similarly, Ontario French has a discourse conjunct with four variants - donc, alors, *ça fait que* and so - which resist ordering (Mougeon et al 1985).

It is fair to say that users of the weighted-index method have generally avoided pitfalls associated with these difficulties. Nevertheless, it is clear that we are still in need of a generalized procedure which is flexible enough to accomodate all the data types found in sociolinguistic inquiry.

Various solutions to the problem have been proposed, and these fall into two groups. The first approach is to analyze the variants separately. Hudson (1980) and Romaine (1980) suggest that a percentage score be calculated for each variant so that particular variants can be associated with particular groups of speakers. While this avoids imposing an order on the variants, it has an important statistical drawback. Any test of significance made on the variants is subject to experimentwise error (that is, there is a notable increase in type I error).

A different avenue is a unitary approach which analyzes the variants simultaneously. Milroy (1982) proposes the creation of two scales: one scale corresponds to the weighted-index, and the second is a range score which measures dispersion among variants for individual subjects or groups of subjects. Other linguists make use of multivariate statistical techniques. Jones-Sargent (1983) uses cluster analysis while Berdan (1978) works with principal components analysis. These techniques study all variants as part of one analysis. The VARBRUL programme (Sankoff 1987) considers multiple variants in a somewhat different light: it may be that the variants are not generated simultaneously but in terms of a sequence of ordered rules each dealing with a binary choice. All possible schemes are considered and the best one is selected using maximum likelihoods. Again, the schemes analyze all variants together. The Correspondence Analysis applications that we present below follow this unitary approach in dealing with multiple variants.

Correspondence analysis

Correspondence analysis is a member of the growing family of multivariate statistical techniques. It is known by different names in different parts of the world - homogeneity analysis, optimal scaling, Hayashi's quantification (see Tenenhaus & Young 1986 for a review); in this paper we will refer to it as Dual Scaling (DS), which is the Canadian version developed by Nishisato (1980). The technique makes no assumptions about data distribution (such as normality) or about the number of observations in each cell of a study, and it is adapted to quantify categorical type data (i.e. categories which cannot be ordered along a continuum). These features make it attractive for exploring the data structures found in sociolinguistic research, and a few applications are found in the literature (for example Cichocki 1986; Inoue 1986). In this paper we focus on two applications in the study of vowel variation.

Glasgow English (i)

This variable is found in words such as *hit*, *thing* and *children*, and has five variants which Macaulay (1978) arranges along a linear scale as follows:

(i - 1) [I]
(i - 2) [I˖] [ɛˆ]
(i - 3) [I˕] [ɛˀ]
(i - 4) [əˆ]
(i - 5) [ʌˆ]

In phonetic terms the variants fall into two dimensions - height and front-backness. Using the weighted-index approach Macaulay finds a linear relation between social group and vocalic variant: speakers from the higher end of the social scale tend to use the higher, more front variants; those from the lower end the mid central variants. His sample includes 16 subjects, four from each socio-economic class (professional, white collar, blue collar, semi-skilled); both genders are equally represented across the social groups.

The DS formulation for this analysis is in terms of a table of frequencies (as in Table 1). The rows are the 16 individual subjects; columns are the 5 variants of (i); the entries are the frequencies with which each subject uses each variant. The procedure treats the rows and columns as two separate spaces - subject space and phonetic space - and objectively determines the optimal association or correlation between them. This association is displayed in the form of a graph which summarizes both indices assigned to variants and scores attributed to subjects. These indices and scores depend strictly on the data structure itself as no preset order is input, and they indicate the extent to which subjects' linguistic performances are similar or different.

The application of DS to the Glasgow English (i) data produces a two-dimensional plane (see Figure 1, p. 46) which accounts for over 90% of the variance in the data. The first (vertical) dimension is the more important (77.2%), the second (horizontal) is less important (16.2%). Locating the phonetic space, we note that the configuration of the vocalic variants corresponds to the familiar pattern of the vowel triangle: the high vowels are at one end of the graph; the mid vowels are spread out along the opposite end. The positions of individual subjects within the phonetic space also present a fairly clear pattern. Members of the highest social group and two females of the second highest group use the high variants; other subjects use the mid variants, showing variation on the front-back phonetic plane.

The most obvious feature of this analysis is the close correspondence it has with Macaulay's (weighted-index) analysis. The order of variants along dimension 1 is exactly the same as their order in his approach. Also, the subject scores calculated by the two approaches are strongly related ($r = -0.988$). This similarity, however, is apparent if one considers the status of the assumption of order among the variants. DS makes no assumptions and arrives at an analysis which is identical to the weighted-index analysis, which requires the ordering assumption.

A second feature is the possibility of examining information represented on the higher dimensions. In particular, it is interesting to examine usage by those subjects who have (near) identical scores on dimension 1 and different scores on dimension 2. Comparing subjects 1 and 7 (in Table 1, p. 45) we observe that subject 1 (professional and male) has a broader range of variants - 30% (i-1), 50% (i-2), 17% (i-3), 3% (i-4) - than subject 7 (white collar and female) - 21% (i-1), 75% (i-2), 4% (i-3). A similar comparison holds for subjects 5 and 14 as well as for subjects 6 and 16. Thus dimension 2 isolates subjects who have a greater (to the right) or lesser (to the left) spread of variants. While this difference is statistically significant, its sociolinguistic relevance and interpretation must be determined by the researcher. Is the spread due to the particular lexical items elicited? Are there attitudinal reasons why certain subjects have greater dispersion of variants than others? Answers to these and similar questions will determine whether or not to retain this higher dimension in the analysis.

A third feature of the DS analysis is the graphic display of the results. Not only does the graph show the data by simultaneously representing subject and phonetic spaces, it also induces the researcher to think further about certain details which may have substantive value for the analysis. Such a visual approach to exploring data is receiving increasing attention among statisticians (Tukey 1977).

This Glasgow English (i) application is an example of the elegant quantification of sociolinguistic variation that DS offers, and it provides some justification for using DS in quantifying variables with multiple variants[2]. We turn now to the analysis of a more complex vocalic variable.

Toronto English (aw)

Toronto English (aw) has six vocalic variants - three low [æw, aw, ɑw] and three mid [ɛw, ʌw, ɒw]- which fit the phonetic dimensions of height and front-backness. The diphthong is found in words such as *mouse*, *trout* and *couch* (in what is called the voiceless environment) as well as in *cow*, *down* and *aloud* (the elsewhere environment). Using data from 18 subjects of three age groups (children, young adults, adults) and both genders, Chambers (1980) shows that there is a (putative) change of (aw)-fronting taking place. The change is occurring regularly across the age groups and is faster in the elsewhere environment. Chambers also discusses a second process, (aw)-non-raising which involves the voiceless environment but which has less well-defined social correlates.

Chambers' analysis is based on a grouping of variants according to the dimensions of vowel space: variant onsets are front [æ, ɛ], central [a, ɒ] or back [ɑ, ʌ] in the case of (aw)-fronting; they are mid [ɛ, ɒ, ʌ] or low [æ, a, ɑ] for (aw)-non-raising. In terms of the problem of quantification, the researcher must be careful here to verify the pattern of the social-linguistic correlation for each of the variants that have been grouped together. This is also possible in a unitary approach which handles all six variants simultaneously. The DS formulation is to combine the two (18 x 6) tables which summarize the usage by the 18 subjects of the 6 variants for the two phonological environments (or word classes). This (36 x 6) table (see Table 2, p. 47) is a super-matrix which we can represent formally as Z = [Zv,Ze], where Zv is the table summarizing data in the voiceless environment and Ze is the table for the elsewhere environment.

The DS analysis of this super-matrix reveals three interpretable dimensions which account for 82.4% of the variance in the array. These are displayed as pairs of planes in Figures 2a, 2b, (p. 48) and 3 (p. 49). Figures 2a and 2b (the dimension 1 - dimension 2 plane) show the vowel triangle with the low vowels spread out at the lower end and the mid vowels grouped together at the upper end. The cluster of mid vowels is broken up in the dimension 1 -dimension 3 plane (Figure 3).

The graph in Figure 2b locates subjects' usage in the elsewhere environment along the [æ]-[a]-[ɑ] continuum. This is the (aw)-fronting sound change, with children leading, followed by young adults and adults. The subject scores calculated by DS correlate very strongly with those of the weighted-index approach (r=-1.00). The same plane (Figure 2a) also represents subjects' production in the voiceless environment. The graph points to four subjects (8, 9, 11 and 15) who do not always raise the (aw) onset - this is (aw)-non-raising which also involves some fronting as shown in the graph. The correlation between the subject scores from the two approaches is strong (r=0.994).

[2]Berdan (1978) proposes a similar approach to the Glasgow data using principal component analysis. Correspondence Analysis is a variant of principal component analysis devised to analyse qualitative data such as the frequency tables which represent this and other variables, while the version used by Berdan is intended for interval and ratio data types. This suggests that data structures derived by the two techniques are not always identical (Lebart, Morineau & Warwick 1984, Chapter 2).

TABLE 1

Frequency-table data matrix for Glasgow English (i)
The table lists frequencies of occurrence of each variant for each subject.

SUBJECT	PHONETIC VARIANTS						
	[I]	[ɛ˄]	[I˅]	[ɛ˃]	[I˳]	[ə˄]	[ʌ˄]
1	24	40		13	3	0	
2	22	48		10	0	0	
3	28	45		7	0	0	
4	47	30		2	0	0	
5	0	16		24	2	0	
6	0	32		37	11	0	
7	10	36		2	0	0	
8	1	38		9	2	0	
9	0	12		36	19	7	
10	1	28		37	13	1	
11	0	22		32	8	0	
12	2	33		38	7	0	
13	0	15		38	22	5	
14	0	34		30	13	3	
15	1	19		21	2	1	
16	1	18		15	5	1	

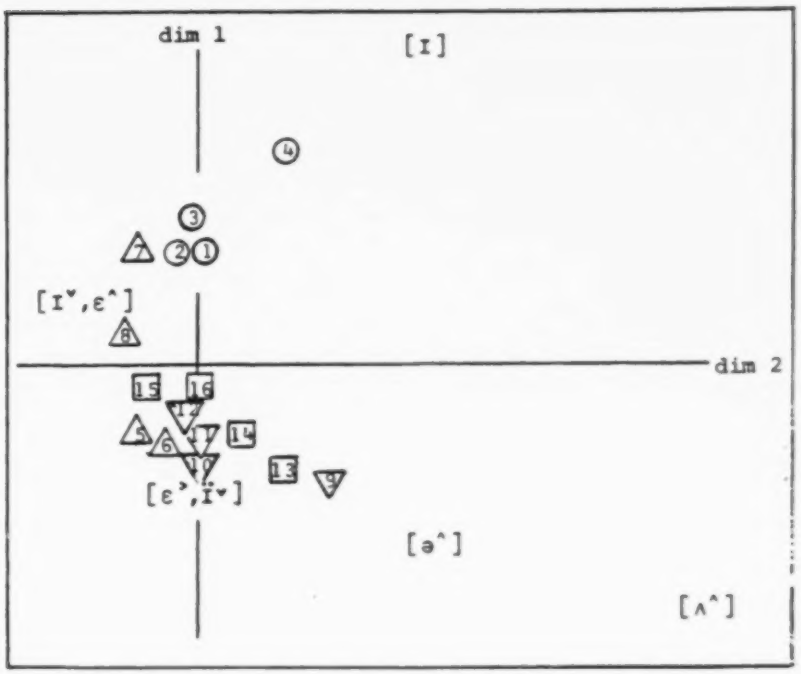
(Source - Macaulay 1978: 140)

The dimension 1 - dimension 3 plane (Figure 3) further analyzes the voiceless environment, focusing primarily on the continuum between [ɒ] and [ʌ]. (We note that the production of [ɛ] is considered to be a separate process and appears on dimension 4.) The process of (aw)-fronting with raising in this plane is characteristic of children, especially boys, who are ahead of girls, and a few male adults. There is a fairly strong correlation with weighted-index scores (r=0.687). This dimension points to a small discrepancy between DS and Chambers' treatments. DS insists that (aw)-fronting is two separate processes depending on the environment, and not one process as in Chambers' analysis. The DS position appears to be due to a gender x environment x process interaction. Females in general lead in the elsewhere environment, which involves (aw)-fronting; males (particularly boys) lead in the voiceless environment, which involves (aw)-fronting with raising. While this difference in linguistic interpretation is not great, neither position overtly distorts the data found in the matrix in Table 2.

The quantification of Toronto English (aw) illustrates that whereas the classical method assumes that variants can be grouped and deals with these groups separately, DS takes a unitary approach. The DS results correspond closely to the classical treatment as shown by the strong subject score correlations. They also show a multidimensional structure which isolates several processes and their interactions. These are: (aw)-non-raising with fronting in the voiceless environment (dimension 1), (aw)-fronting in the elsewhere environment (dimension 2), (aw)-fronting with raising in the voiceless environment (dimension 3), and the production of [ɛ] (dimension 4). For the most part these receive a reasonable linguistic interpretation.

FIGURE 1:

Graph of the dimension 1 - dimension 2 plane for Glasgow English (i)
Five vocalic variants constitute phonetic space.
Each number locates a subject in phonetic space;
social group membership is marked as:
○ professional, △ white collar, ▽ skilled, □ semi-skilled



Conclusion

This paper shows how Correspondence Analysis provides an objective and generalized approach to the quantification of variables with more than two variants. The technique makes no assumptions about the distribution of sociolinguistic data such as the order of variants, and it uses a mathematically-based approach to produce results which are linguistically interpretable and which correspond closely with those proposed in studies using the weighted-index method.

Among the advantages of the technique are its unitary approach (which coincides with statistical practice and which avoids experimentwise error), its use of multidimensionality (so that significant and interesting patterns are isolated beyond those one finds in a linear or unidimensional perspective), and its graphical representation (which both clarifies the description of variation and enhances our understanding of it). Of course, the linguistic interpretability of the solutions extracted will ultimately determine the usefulness of the technique, as will further empirical testing on other data.

TABLE 2:

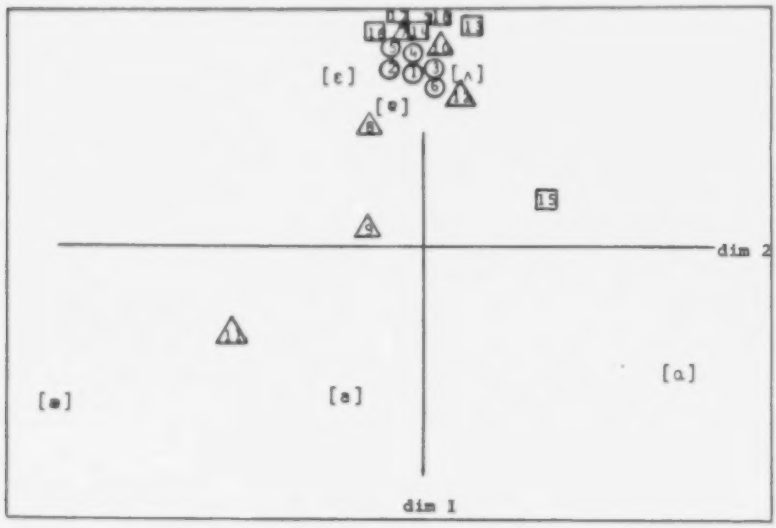
Frequency-table data matrix for Toronto English (aw)
The super-matrix lists frequencies of occurrence of each variant for each subject in both voiceless and elsewhere environments

SUBJECT	PHONETIC VARIANTS [ɑ]	[ʌ]	[ə]	[ɒ]	[ɐ]	[ɛ]	
1	0	7	0	8	0	0	
2	0	3	0	12	0	0	
3	1	6	0	8	0	0	
4	0	9	0	6	0	0	
5	0	9	0	5	0	1	
6	1	10	0	3	0	0	
7	0	15	0	0	0	0	
8	0	11	1	0	1	2	
9	0	0	8	7	0	0	voiceless
10	0	15	0	0	0	0	environment
11	0	2	7	0	6	0	
12	1	13	1	0	0	0	
13	0	14	0	1	0	0	
14	0	14	0	1	0	0	
15	7	8	0	0	0	0	
16	0	15	0	0	0	0	
17	0	12	0	3	0	0	
18	0	14	0	1	0	0	
1	4	0	11	0	0	0	
2	1	0	14	0	0	0	
3	5	0	7	1	2	0	
4	5	0	4	0	6	0	
5	0	0	4	0	11	0	
6	0	0	3	0	12	0	
7	7	0	8	0	0	0	
8	6	0	6	0	3	0	
9	0	0	14	0	1	0	elsewhere
10	9	0	6	0	0	0	environment
11	1	0	8	0	6	0	
12	5	0	9	0	1	0	
13	10	0	5	0	0	0	
14	13	0	2	0	0	0	
15	13	0	1	0	0	0	
16	15	0	0	0	0	0	
17	5	0	10	0	0	0	
18	15	0	0	0	0	0	

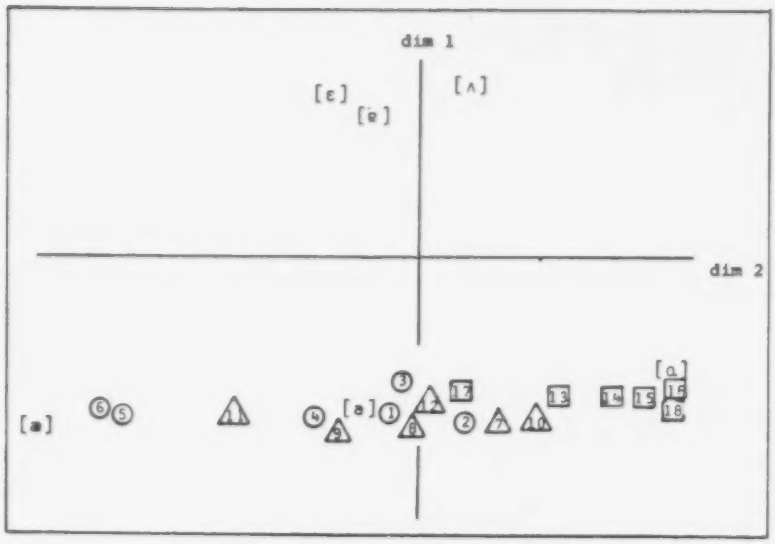
(Source - Chambers p.c.)

FIGURE 2:

Graphs of the dimension 1 - dimension 2 plane for Toronto English (aw), based on DS analysis of data super-matrix. Subjects are located in phonetic space according to their production in (a) voiceless and (b) elsewhere environments. Age group membership is marked as: ○ children, △ young adults, □ adults.



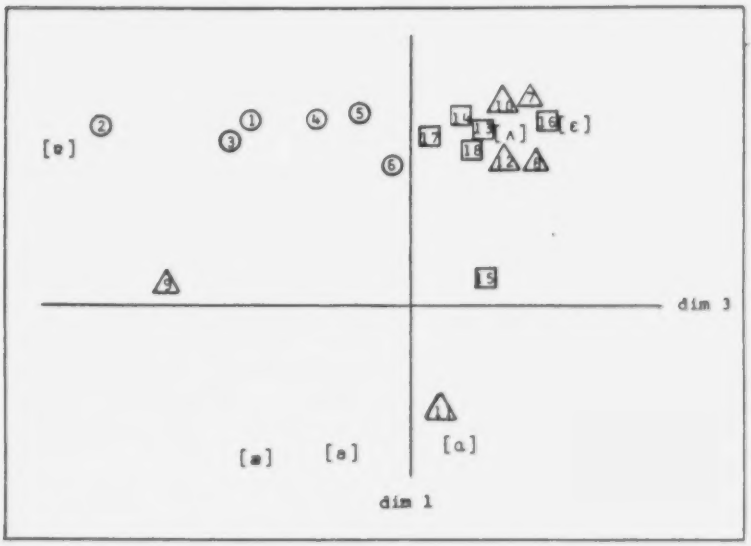
(a) voiceless environment



(b) elsewhere environment

FIGURE 3:

Graph of the dimension 1 - dimension 3 plane for Toronto English (aw)
Each number indicates the position of a subject in phonetic space,
given his/her production in the voiceless environment.
Age group membership is marked as: ○children, △young adults, □adults.



BIBLIOGRAPHY

BERDAN, R. (1978), "Multidimensional analysis of vowel variation", in D.Sankoff (ed), *Linguistic variation: models and methods*, pp. 149-60.

CHAMBERS, J.K. (1980), "Linguistic variation and Chomsky's 'homogenous speech community'", in A.M. Kinloch and A.B. House (eds), *Papers from the fourth meeting of the Atlantic Provinces Linguistic Association*, pp. 1-32.

CICHOCKI, W. (1986), *Linguistic Applications of Dual Scaling in Variation Studies*, unpublished Ph.D. dissertation, University of Toronto.

HUDSON, R. (1980), *Sociolinguistics*, Cambridge University Press, Cambridge, UK.

INOUE, F. (1986), "Sociolinguistic aspects of new dialect forms: language change in progress in Tokyo", in *International Journal of the Sociology of Language*, 58, pp. 73-89.

JONES-SARGENT, V. (1983), *Tyne Bytes: A computerized Sociolinguistic Study of Tyneside*, Verlag Peter Lang, New York.

KEMP, W. (1979), "L'histoire récente de CE QUE, QU'EST-CE QUE et QU'OSQUE à Montréal: Trois variantes en interaction", in P. Thibault (ed), *Le français parlé: Études sociolinguistiques*, Linguistic Research, Edmonton, pp. 53-74.

LABOV, W. (1966), *The social stratification of English in New York City*, Center for Applied Linguistics, Washington D.C.

LEBART, L., MORINEAU, A. & WARWICK, K.M. (1984), *Multivariate descriptive statistical analysis*, John Wiley & Sons, New York.

MACAULAY, R. (1978), "Variation and consistency in Glasgow English", in P.Trudgill (ed), *Sociolinguistic patterns of British English*, Edward Arnold, London, pp. 132-43.

MILROY, J. (1982), "Probing under the tip of the iceberg: phonological 'normalization' and the shape of speech communities", in S. Romaine (ed), *Sociolinguistic variation in speech communities*, Edward Arnold, London, pp. 35-48.

MOUGEON, R., BENIAK, E. & VALOIS, D. (1985), "Issues in the study of language contact: evidence from Ontarian French", unpublished MS, Ontario Institute for Studies in Education.

NISHISATO, S. (1980), *Analysis of categorical data: Dual Scaling and its applications*, University of Toronto Press, Toronto.

ROMAINE, S. (1978), "Postvocalic /r/ in Scottish English: sound change in progress?", in P. Trudgill (ed), *Sociolinguistic patterns in British English*, Edward Arnold, London, pp. 144-58.

ROMAINE, S. (1980), "A critical overview of the methodology of urban British sociolinguistics", in *English World Wide*, 1/2, 163-98.

ROMAINE, S. (ed) (1982), *Sociolinguistic variation in speech communities*, Edward Arnold, London.

SANKOFF, D. (ed) (1978), *Linguistic variation: models and methods*, Academic Press, New York.

SANKOFF, D. (sous presse), "Variable rules", to appear in U. Ammon, N. Dittmar & K.J. Mattheir (eds) *Sociolinguistics. An international handbook of the science of language and society.*

TENENHAUS, M. & YOUNG, F.W. (1986), "An analysis and synthesis of multiple correspondence analysis, optimal scaling, dual scaling, homogeneity analysis and other methods for quantifying categorical multivariate data", in *Psychometrika*, 50, 91-119.

THIBAULT, P. (ed) (1979), *Le français parlé: Études sociolinguistiques*, Linguistic Research, Edmonton.

TRUDGILL, P. (ed) (1978), *Sociolinguistic patterns in British English*, Edward Arnold, London.

TUKEY, J. (1977), *Exploratory data analysis*, Addison-Wesley, Reading, MA.

UNE ÉTUDE DES QUALIFICATEURS, INTENSIFICATEURS, ET PHRASES ENCLOSES PAR RAPPORT À LA DIFFÉRENCIATION LINGUISTIQUE SELON LE SEXE[1]

Charlene Potter
Université d'Ottawa

À partir de Jespersen (1922) jusqu'à Labov (1972), on traitait le parler des femmes seulement en le comparant au parler des hommes. Cette comparaison était souvent faite au détriment du parler des femmes, et bien sûr, souvent des femmes elles-mêmes. Par exemple, Labov (1972) a dit que les femmes sont "insécures linguistiquement", alors que Shuy (1970) a suggéré que les femmes doivent rester, pour les linguistes, "un des mystères de l'univers". Je n'accepte pas cette façon de penser. C'est désavantageux pour les femmes en ce sens que cette idée les limite et les empêche de grandir, de se développer.

Les données de mon étude consistent en douze heures d'entrevues enregistrées. Les informateurs et les informatrices[2] sont tous/toutes de bon(ne)s ami(e)s et les entrevues sont donc informelles.

J'ai exploré trois possibilités de base: une femme interviewant une autre femme, un homme interviewant un autre homme, et une femme interviewant un homme. Malheureusement, je n'ai pas un homme interviewant une femme, mais j'ai l'intention d'ajouter cette dyade. Chacune de ces trois dyades se compose de deux entrevues.

Au départ, je voulais étudier des "tag questions", par exemple, "You liked the movie, didn't you?" mais il n'y avait pas assez de ces expressions dans mon corpus.

C'est alors que la professeure Poplack m'a aidée à trouver un autre sujet de recherche. Nous avons remarqué qu'il y avait beaucoup d'expressions sans contenu très important; par exemple, l'expression "kind of" dans la phrase: "The book was kind of interesting.". A ce moment-là, j'ai choisi d'étudier les qualificateurs, les intensificateurs et ce qu'on pourrait appeler les "phrases encloses" (en anglais: "hedges"). Une phrase enclose est une phrase qui contient un mot ou une expression qui enlève de la force à l'affirmation de cette phrase; par exemple, l'expression "sort of", dans la phrase "This table is sort of hard", enlève de la force à l'affirmation "This table is hard".

J'ai cherché à savoir s'il y avait une différence de fréquence de ces variables entre le parler des femmes et celui des hommes. Des études précédentes montraient que le parler des femmes était plus qualificateur, moins ferme et moins direct. J'y reviendrai plus loin, mais voici d'abord comment j'ai analysé mes données.

J'ai pris cent phrases de chaque informateur et informatrice. Chaque entrevue se divisait en quatre sections: la famille ou les ami(e)s, le travail, les études et les vacances. J'avais à peu près le

---

[1]Ceci est un travail de recherche que j'ai commencé l'an passé alors que j'étais encore au niveau du baccalauréat. J'ai l'intention de le reprendre et de l'approfondir pour en faire ma thèse de maîtrise.

[2]J'aimerais expliquer que je crois que les femmes sont exclues sous un nom masculin. Je vais utiliser les deux formes; par exemple: informateurs et informatrices. Je vais aussi accorder tous les adjectifs qualificatifs aux deux genres. En faisant cela, on ne fait pas disparaître tout un être dans un discours masculin. Je sais que cela va prendre plus de temps et qu'il y a bien des gens qui n'aiment pas cela mais je trouve que c'est bien nécessaire.

même nombre de phrases dans chacune de ces sections. Parmi ces cent phrases, j'ai compté celles qui contenaient une ou plusieurs variables recherchées comme étant "marquées". Si une phrase ne contenait pas de variable, elle était non-marquée.

Hartman (1976) a dit que le parler des femmes est "fleuri, peu assuré et tempéré" (en anglais: "flowery, tentative and qualified"). Mais les résultats que j'ai obtenus ne montrent pas tout à fait la même chose...

Dans les tableaux, "F" veut dire "femme" et "H" veut dire "homme". Les lettres majuscules dans les colonnes représentent les six informateurs et informatrices et le petit "t" sur la dernière ligne veut dire "total".

TABLEAU 1

Phrases marquées / phrases totales

F-H	F-F	H-H
C 48 / 100 [48%]	M 30 / 100 [30%]	G 51 / 100 [51%]
G 46 / 100 [46%]	T 31 / 100 [31%]	D 53 / 100 [53%]
t 94 / 200 [47%]	t 61 / 200 [30,5%]	t 104 / 200 [52%]

Le tableau numéro 1 montre le nombre de phrases marquées par rapport au nombre total de phrases. On voit que c'est la dyade F-F (femmes interviewant femme) qui contient le plus petit pourcentage de phrases marquées: 30,5%. La dyade F-H (femme interviewant homme) montre 47% de phrases marquées et la dyade avec le plus de phrases marquées c'est H-H (homme interviewant homme) avec une proportion de 52%. Déjà, en général, on voit que c'est le parler des hommes qui contient le plus fort pourcentage de phrases marquées.

Ensuite, j'ai mis toutes mes variables dans les catégories: qualificateurs, intensificateurs et phrases encloses, pour montrer exactement "qui a dit quoi et combien de fois". Ainsi, j'ai éliminé les cas ambigus où des phrases contenaient plus d'une variable.

L'expression "you know" est la première variable que j'ai observée. La variable "you know" marque une phrase enclose. Lorsque le verbe est employé, je ne le compte pas; par exemple: "Do you know the capital of Alberta?". L'expression "you know" compte dans une phrase où sa présence n'est pas nécessaire, c'est-à-dire qu'on peut la laisser tomber et que la phrase sera encore bonne; par exemple: "It was **you know** a nice day".

Pamela Fishman (1980) a observé l'expression "you know". Elle a enregistré des conversations quotidiennes de jeunes couples américains. Elle a trouvé que les femmes disent "you know" cinq fois plus souvent que les hommes. Fishman définit l'expression "you know" comme un moyen d'obtenir l'attention de quelqu'un, par exemple, quand une personne veut vérifier si son interlocuteur la com-

prend. Dans son étude, Fishman explique que les femmes disent "you know" plus souvent parce que les hommes ne les écoutent pas. Il fallait donc aux femmes de ces jeunes couples américains un moyen d'obtenir l'attention des hommes et c'est pourquoi elles utilisent l'expression "you know".

TABLEAU 2

"You know" / total de phrases marquées

F-H	F-F	H-H
C 6 / 48 [12,5%]	M 0 / 30 [0%]	G 4 / 51 [7,8%]
G 5 / 46 [10,9%]	T 0 / 31 [0%]	D 5 / 53 [9,4%]
t 11 / 94 [11,7%]	t 0 / 61 [0%]	t 9 / 104 [8,7%]

Le tableau numéro deux montre la distribution de "you know". Comme on peut le constater, mes résultats sont différents des résultats de Fishman. La dyade F-F ne l'utilise pas du tout et les dyades H-H et F-H l'utilisent à environ 10%.

J'ai ensuite observé d'autres phrases encloses qui contenaient cette fois-ci les expressions "kind of" et "sort of" que j'ai regroupées parce qu'elles sont interchangeables; par exemple: "The movie was **kind of** good" ou "The movie was **sort of** good".

TABLEAU 3

"Sort of" et "kind of"[3] / Total de phrases marquées

F-H	F-F	H-H
C 1 / 48 [2,1%]	M 1 / 30 [3,3%]	G 4 / 51 [7,8%]
G 1 / 46 [2,2%]	T 0 / 31 [0%]	D 19 / 53 [35,8%]
t 2 / 94 [2,1%]	t 1 / 61 [1,6%]	t 23 / 104 [22,1%]

[3]Sont aussi incluses les expressions suivantes: "sort of like" et "kind of like".

En regardant le tableau numéro trois, on peut penser que l'informateur "D" biaise les résultats. Mais même sans l'informateur "D", cette dyade montrerait toujours le plus haut pourcentage. Dans les deux autres dyades (F-F et F-H), on voit que cet emploi est très rare. Mes résultats montrent donc que ce seraient plutôt les hommes dans les tableaux deux et trois qui emploient le plus souvent les qualificateurs.

Ensuite, j'ai observé les expressions comme "I think", "I guess", "I dunno" (I don't know) et "I'm not sure". Selon Coates (1986), la fonction principale d'une expression comme "I think" est d'exprimer l'incertitude et la déférence; utiliser "I think", c'est se soumettre au jugement de l'autre ou c'est rendre la phrase plus polie. Par exemple, si quelqu'un dit: "It's nice out", cela veut dire qu'il fait beau. Mais si quelqu'un dit la phrase: "I think it's nice out", c'est selon Coates une phrase moins puissante et plus polie, tout comme le parler des femmes.

Perkins (1983) a trouvé la même tendance chez les enfants. Il a trouvé que les filles disaient "I think" beaucoup plus souvent que les garçons (61,3% versus 38,7%). Perkins suggère que sa découverte appuie l'idée que les filles apprennent moins à s'affirmer en acquérant la compétence communicative.

Le tableau numéro quatre montre mes résultats pour ces expressions d'incertitude et de déférence.

TABLEAU 4

"I think", "I guess", "I dunno", "I'm not sure", etc. / total de phrases marquées

F-H	F-F	H-H
C 4	M 0	G 7
48 [8,3%]	30 [0%]	51 [13,7%]
G 12	T 2	D 7
46 [26,1%]	31 [6,5%]	53 [13,2%]
t 16	t 2	t 14
94 [17,0%]	61 [3,3%]	104 [13,5%]

On remarque que les femmes utilisent à une fréquence de seulement 3,3% ces expressions alors que les hommes les utilisent à plus de 13%. Il est donc évident qu'il y a une différence et que ce ne sont pas les informatrices qui montrent de l'incertitude et de la déférence.

La catégorie suivante était celle des intensificateurs; il s'agit de mots comme "so", "such", "very", "really" et "quite". Jespersen (1922) a dit que "ce sont les intensificateurs qui sont responsables du manque de précision du parler des femmes".

Lakoff (1975) a dit que les femmes utilisent le mot "so" plus souvent que les hommes. Puis Key (1975) a dit que les femmes utilisent le mot "such" plus souvent que les hommes.

TABLEAU 5

Intensificateurs: "so", "such", "very", "really", "real", "quite" / Total de phrases marquées

F-H	F-F	H-H
C 12 / 48 [25,0%]	M 10 / 30 [33,3%]	G 12 / 51 [23,5%]
G 15 / 46 [32,6%]	T 10 / 31 [32,3%]	D 13 / 53 [24,5%]
t 27 / 94 [28,7%]	t 20 / 61 [32,8%]	t 25 / 104 [24,0%]

En regardant les données du tableau cinq, on constate, en effet, que les femmes utilisent plus d'intensificateurs que les hommes mais que la différence n'est pas très grande. Donc, je ne me risquerais pas à tirer une conclusion aussi définitive que celles de Jespersen, Lakoff ou Key.

Nous sommes rendu(e)s à la variable "well". Cette variable ne compte pas quand elle sert à comparer deux choses; par exemple: "I will study as well as go to the meeting". Elle ne compte pas non plus quand elle se trouve au début d'une phrase. Je ne l'ai pas comptée au début des phrases parce que souvent, quand on répond à une question, on commence à réfléchir et, en pensant, on dit "well". Ainsi le nombre de "well" aurait été biaisé selon le nombre de questions posées. Par contre, "well" compte dans une phrase comme celle-ci: "He starts pointing it at us well especially at Larry". Dans cette phrase "well" diminue l'affirmation et rend la phrase moins directe.

TABLEAU 6

"Well" / Total de phrases marquées

F-H	F-F	H-H
C 5 / 48 [10,4%]	M 1 / 30 [3,3%]	G 4 / 51 [7,8%]
G 5 / 46 [10,9%]	T 1 / 31 [3,2%]	D 3 / 53 [5,7%]
t 10 / 94 [10,6%]	t 2 / 61 [3,3%]	t 7 / 104 [6,7%]

En regardant le tableau numéro six, on constate, encore une fois, que les femmes ont le plus bas pourcentage (3,3%) et que les hommes ont le plus haut pourcentage (6,7%). Donc je ne peux appuyer Lakoff (1975) quand elle dit que les femmes disent "well" plus souvent parce qu'elles sont moins sûres d'elles-mêmes.

La dernière variable que j'ai observée est le mot "like". Je ne le compte pas quand il est employé comme verbe; par exemple: "I like you", ni comme comparatif; par exemple: "Do it like this, not like that". Je le compte quand il a perdu son contenu sémantique. Par exemple, un "remplisseur" qui rend la phrase moins directe et moins ferme, comme dans les phrases suivantes:

"So he like supervises the buiding."
"It's like we didn't want to see it."
"I like failed the exam by six marks."

Théoriquement, on s'attend à ce que les femmes utilisent plus souvent ce mot-là. Mais comme on voit dans le tableau numéro sept, c'est tout le contraire.

TABLEAU 7

"Like" / Total de phrases marquées

	F-H	F-F	H-H
C	10 48 [20,8%]	M 0 30 [0%]	G 14 51 [27,5%]
G	14 46 [30,4%]	T 0 31 [0%]	D 10 53 [18,9%]
t	24 94 [25,5%]	t 0 61 [0%]	t 24 104 [23,1%]

Le parler des femmes ne contient pas du tout le mot "like". Dans le parler des hommes, le niveau est assez élevé (H-H: 23,1% et F-H: 25,5%).

Pour tirer une conclusion, j'ai reporté les résultats dans le tableau numéro huit.

TABLEAU 8

Les résultats

TABLEAU	VARIABLE	plus petit %	moyen %	plus grand %
1	Total des phrases marquées	F-F	F-H	H-H
2	"You know"	F-F	H-H	F-H
3	"sort of" et "kind of"	F-F	F-H	H-H
4	"I think", "I guess" ...	F-F	F-H	H-H
5	"so", "such", "very" ...	H-H	F-H	F-F
6	"well"	F-F	F-H	H-H
7	"like"	F-F	H-H	F-H

Dans la colonne du plus petit pourcentage, on voit que c'est le parler des femmes qui montre le moins de phrases marquées; à une exception près. En retournant au tableau numéro cinq, on voit cette exception et on remarque comme auparavant qu'il n'y a pas beaucoup de différences entre ces pourcentages.

Donc, cette étude pilote ne fait pas ressortir les mêmes tendances langagières des parlers des femmes et des hommes que les études précédentes de Lakoff, Key, Perkins et Hartman, etc.

Leurs études disaient que le parler des femmes est fleuri, moins assuré et moins direct. Mon étude montre le contraire. C'est ici que s'arrête ce travail de recherche et j'ai bien hâte de voir si le travail plus approfondi que je ferai va donner les mêmes résultats.

BIBLIOGRAPHIE

CAMERON, Deborah (1985), *Feminism and linguistic theory*, MacMillan Press.

COATES, Jennifer (1986), *Women, men, and language: A sociolinguistic account of sex differences in language*, Longman.

FISHMAN, Pamela M. (1980), "Conversational insecurity", dans *language: Social psychological perspectives*, Pergammon Press.

JESPERSON, Otto (1922), *Language: Its nature, development and origin*, Allen and Unwin.

HARTMAN, Maryann (1976), "A descriptive study of the language of men and women born in Maine around 1900 as it reflects the Lakoff hypothesis in *Language in women's place*", dans *The sociology of languages of American women*, P.I.S.E Papers, IV.

KRAMARAE, Cheris (1981), *Women and men speaking: Framework for analysis*, Newbury House.

KEY, Mary Ritchie (1975), *Male-female language*, Scarecrow Press.

LABOV, William (1972), *Sociolinguistic patterns*, University of Pennsylvania Press.

LAKOFF, Robin (1975), *Language and women's place*, Harper and Rowe.

PERKINS, M. (1983), *Modal expressions in English*, Frances Pinter.

SPENDER, Dale (1980), *Man-made language*, Routeledge and Keegan Paul.

THORNE, BARRIE ET AL. (eds) (1983), *Language, gender and society*, Newbury House.

REMARQUES

J'aimerais remercier de leur aide:

Pierre C. Carbonneau
Pierrette Carrière
Christopher Church
Geraldine Finn
Shana Poplack

Merci aussi à tout le monde qui a participé à cette étude et à ceux qui ont participé au colloque des 25 et 26 septembre à l'Université Laval.

ORGANISATION CONVERSATIONNELLE EN SITUATION DE PLURILINGUISME: LES CHOIX LINGUISTIQUES DES ITALOPHONES DE MONTRÉAL[1]

Normand Labrie
Université Laval

En sociolinguistique, l'étude de la variation s'est faite depuis ses débuts sur deux plans majeurs: la distribution et l'emploi de certaines variables à l'intérieur d'une même variété, ce qui représente l'objet propre à cette discipline et, de façon complémentaire, la distribution et l'emploi de différentes variétés dans un même espace linguistique. Ce travail portera sur une étude menée en situation de plurilinguisme. Elle concernera donc plus précisément ce second champ d'investigation. Suite à un examen en survol du corpus recueilli auprès d'un groupe d'italophones de Montréal qui font usage de l'italien, du français et de l'anglais, nous présenterons d'abord un tableau général de leur utilisation des langues. Puis nous tenterons de dresser un inventaire le plus complet possible des différentes manifestations langagières liées à la "cohabitation" linguistique.

Méthodologie

L'enquête a été menée en 1985 et 1986 dans 16 familles de la partie est de l'Ile de Montréal. Deux locuteurs par famille ont fait l'objet d'une analyse détaillée au plan de leurs comportements linguistiques, soit le père de famille, immigrant de l'après-guerre et l'un de ses fils actif sur le marché du travail au moment de l'enquête. Quarante-cinq heures d'enregistrement ont été produites auprès de 31 locuteurs en situation naturelle. L'ensemble des locuteurs ont été enregistrés lors d'un repas en famille en l'absence de l'enquêteur. Puis, un sous-échantillon formé des volontaires (soit 14 locuteurs) a été enregistré sur les lieux du travail.

Les enregistrements ont d'abord fait l'objet d'une description détaillée en fonction de la totalité des tours de parole produits dans chaque langue. La première partie de cet exposé, portant sur les choix linguistiques des italophones de Montréal, s'appuiera sur ces données.

Les enregistrements ont par ailleurs fait l'objet d'une transcription littérale sélective, c'est-à-dire que seules les séquences contenant des manifestations du contact des langues italienne, française ou anglaise ont été retenues pour la transcription. Les passages entièrement unilingues ont été mis de côté. Il en a résulté 400 pages de transcriptions comprenant près de 5000 manifestations du contact des langues. La seconde partie de l'exposé sur les comportements linguistiques s'appuiera sur l'examen préliminaire de ces données.

Les choix linguistiques des italophones

Au total, nous avons répertorié 39 994 tours de parole produits par l'ensemble des personnes concernées par les enregistrements. De ce nombre, 18 681 sont prononcés en italien (soit 46,7%); 11 903 en français (soit 29,8%); 9 404 en anglais (soit 23,5%); et seulement 6 dans une autre langue (soit 0%). Un nouveau tour de parole débute en moyenne toutes les 4,06 secondes.

---

[1]La présente recherche est menée avec l'aide d'une bourse de doctorat du CRSHC. La tranacription du corpus a été rendue possible grâce à une subvention de l'Office de la langue française. Je remercie ma directrice de thèse Denise Deshaies et le professeur Shana Poplack pour leurs conseils judicieux ainsi que les assistants qui ont travaillé à la transcription du corpus, Raimondo Bussanga, Chiara Morello-Letendre et Columba Inneo.

Les locuteurs dont les comportements font l'objet de l'analyse plus détaillée au sujet des manifestations du contact des langues ont produit au total 15 423 tours de parole, soit 38,6% de l'ensemble des tours de parole. Les locuteurs sur lesquels porte plus exclusivement cette étude ont donc pris une part active à la conversation sans toutefois avoir fait preuve d'une monopolisation de la parole en tant qu'enquêtés. La proportion d'emploi des trois langues par les locuteurs étudiés est sensiblement identique à celle qui a été observée précédemment pour l'ensemble du corpus (soit pour l'italien 50,4%, pour le français 29,8% et pour l'anglais 19,8%).

Ces données très générales permettent de situer les manifestations du contact des langues dans le contexte plus large de l'ensemble du discours. Ainsi, si l'on considère l'ensemble des occurrences produites par les locuteurs visés par l'étude (4848 occurrences), c'est-à-dire si l'on tient compte autant des changements de langue que des alternances ou des intégrations lexicales, on doit reconnaître que près du tiers (31,4%) de tous les tours de parole produits par les locuteurs qui nous intéressent sont liés au phénomène du contact des langues: *le phénomène du contact des langues touche environ un tour de parole sur trois.*

Examinons de plus près en quoi consistent ces manifestations du contact des langues[2]. Nous allons voir comment se répartissent les options linguistiques (voir la Figure en annexe: Distribution des occurrences impliquant deux langues[3]) et dans quelle mesure les locuteurs y ont recours en fonction des différentes combinaisons de langues (voir le Tableau en annexe: Le corpus des italophones de Montréal).

Les grandes catégories d'analyse

Le premier grand niveau d'analyse est de type conversationnel. En prenant en compte les changements de langue à la prise de parole, on distingue trois types de comportements distincts.

1) Celui où le locuteur diverge de la langue qui vient d'être utilisée par son interlocuteur en restant fidèle cependant à la langue dans laquelle il avait produit son dernier tour de parole (DC) (ex. 1).

Exemple 1:

0003 A: buono il brodo
0003 G: you like it
0003 A: * ti piace il formaggio dentro

Ce type de comportement est de loin le plus utilisé par les locuteurs étudiés puisqu'il concerne 1596 occurrences, soit 32,9% de l'ensemble du corpus ou le tiers de toutes les manifestations du

[2]Le système de classification des 4848 occurrences à analyser a été établi sur la base d'un examen approfondi du corpus ainsi que sur la base de travaux connus en analyse conversationnelle ou en linguistique. Les principaux ouvrages qui ont servi sont les suivants : Auer (1984), Gumperz (1982), Poplack (1980), Poplack et Sankoff (sous presse), Poplack, Sankoff et Miller (1987), Sankoff (1980). Le modèle d'analyse a été décrit plus en détail dans un autre article (Labrie, sous presse).

[3]Les sigles qui identifient les types d'occurrences (DC, CD, DD, EP, IP, Mot, Nom) sont expliqués plus bas. Les langues impliquées sont identifiées comme suit : IF, de l'italien vers le français; IA, de l'italien vers l'anglais, etc.

contact des langues rencontrées. Il s'agit en fait d'un type de comportement qui favorise le maintien de la langue. Cette divergence par rapport à l'interlocuteur repose sur l'imposition aux autres d'une langue utilisée de façon constante.

2) Celui où le locuteur converge vers la langue qui vient d'être utilisée par son interlocuteur en abandonnant la langue dans laquelle il avait lui-même produit son dernier tour de parole (CD) (ex. 2).

Exemple 2:

0001 A: quante quanta carne fai ?
0001 G: quelle four chops che son fuori lì (.) and a glass of wine
0001 please (0.03) honey ?
0001 A: * there you come again =

Ce type concerne 824 occurrences ou 17,0% de toutes les occurrences du corpus. Il s'agit d'un comportement d'adaptation. Cette convergence vers l'interlocuteur implique une certaine souplesse linguistique de la part du locuteur.

3) Et enfin, celui où le locuteur diverge à la fois de la langue utilisée tant par son interlocuteur que par lui-même lors de son dernier tour de parole (DD) (ex. 3).

Exemple 3:

0004 G: I don't know credo che non l'avevi già messo so what are you
0004 going to do tonight ?
0004 A: * nothing
0005 G: you're not going out or anything ?
0005 A: * devo andare da Renaldo a portargli l'assegno

Ce type de comportement touche 731 occurrences dont 119 impliquent l'utilisation consécutive de trois langues. La répartition plutôt égale de ces 119 occurrences trilingues est donnée dans la Figure en annexe: Distribution des 119 occurrences impliquant 3 langues. Au total, les 731 occurrences de ce type représentent 15,1% de toutes les occurrences du corpus. Il s'agit de la forme la plus dynamique dans la distribution des langues en conversation.

Si les trois catégories conversationnelles décrites ci-haut correspondent à des catégories factuelles qui font abstaction des motivations du locuteur, elles donnent néanmoins une mesure de son activité linguistique. Ces trois types d'alternance conversationnelle font partie de la mécanique conversationnelle. Ce sont les catégories les plus pertinentes du point de vue de la fréquence de leur utilisation, car au total on retrouve 3151 occurrences qui sont regroupables de la sorte, ce qui équivaut à 65,0% des toutes les manifestations du contact des langues que l'on retrouve dans le corpus. Alors que cette majeure partie du corpus implique des changements de langue à la charnière des tours de parole, le reste des occurrences, soit 35,0% du corpus, est réparti entre divers phénomènes internes au tour de parole. Il s'agit de l'alternance entre deux phrases, de l'alternance entre des éléments d'une même phrase et de l'emploi de mots isolés.

Le prochain niveau d'analyse est de type intermédiaire puisqu'il se produit à l'intérieur d'un même tour de parole sans toutefois mettre en jeu la structure syntaxique de la phrase. Il s'agit de l'alternance extra-phrastique, l'alternance entre deux phrases. On peut distinguer deux genres princi-

paux de comportements selon que l'alternance extra-phrastique est liée au fait que le locuteur s'adresse à un nouvel interlocuteur (type conversationnel, ex. 4) ou qu'elle est produite sans qu'il y ait changement d'interlocuteur (type linguistique, ex. 5).

Exemple 4:

0526 L: y a venu ici le lendemain y était fâché tu t'sais * remember ?
0526 he was telling us his story (XXX) (0.08)

Exemple 5:

0066 G: if it's their second home and they have grown up kids
0066 [seventeen eighteen]
0066 A: * [it's two more downstairs] =
0067 = * ne han /hanno/ due già

Le type conversationnel de l'alternance extra-phrastique concerne seulement 79 occurrences, soit 1,6% de toutes les occurrences que compte le corpus. Le type linguistique d'alternance extra-phrastique comprend surtout des alternances entre deux phrases qui s'adressent à un même interlocuteur. On compte 283 occurrences du genre ou 5,8% du corpus. En outre, il existe d'autres genres d'alternances extra-phrastiques, comme par exemple, des occurrences de discours rapporté sans identification (16 occurrences, ex. 6) et des faux-départs, où l'on a une première phrase tronquée, suivie d'une phrase entièrement nouvelle (le type p1- p2, 18 occurrences, ex. 7), ou encore une première phrase tronquée, suivie d'une nouvelle phrase dans laquelle il y a reprise d'un SN sujet équivalent (le type p1- p1, 8 occurrences, ex. 8).

Exemple 6:

0543 L: ..."t'as rien que ben" j'y dis "ben (XXX) en train de
0543 s'embrasser et tout chus pas pour aller là puis y dire ça" y
0543 dit "aie y faut" y dit y dit "y a tellement d'espace dans
0543 l'parc y doivent pas venir ici ()"(XXX) * "it's all right" * (XXX)

Exemple 7:

0266 R: ma non cono- e. le- * je l'sais pas j'espère que ça l'dérange

Exemple 8:

0662 L: alors ce gars cet homme joueur de football habillé en femme
0662 c'était le body-guard de la mère =
0663 = (XXX) alors y avait un gars qui voulait attaquer la mère
0663 j'sais pas quoi le gars y le * he tackled him =

Au total donc, 404 occurrences sont de type extra-phrastique, soit 8,3% des occurrences du corpus. Cette catégorie fait davantage appel à une interprétation linguistique que conversationnelle. Elle sous-tend en fait chez le locuteur des compétences dans chacune des langues impliquées par l'alternance.

Le prochain niveau d'analyse est de type intra-phrastique. On compte dans cette catégorie des formes périphériques telles que l'utilisation d'interjections en début d'énoncé (94 occurrences, ex. 9), l'utilisation de syntagmes libres à l'intérieur d'un énoncé (81 occurrences, ex. 10), des occurrences de discours rapporté identifiées (32 occurrences, ex. 11), des appositions (traduction-répétition/reprise, non reprise, expression/explicitation) (25 occurrences, ex. 12), des subordinations (21 occurrences, ex. 13), des coordinations (13 occurrences, ex. 14).

Exemple 9:

0426 X: fa un comico che canta che fa (0.02) e. prestidigitazione un
0426 affare cosi che fantesista /fantasista/ no Linda de Susa (.)
0426 sarebbe una portoghese che hanno scoperto in Francia
0426 M: * oua oua oua oui * sta l'ho vista alla televisione che

Exemple 10:

2236 M: viene sen- sentimmo tre quattre lingue alla volta * tabarnacle*

Exemple 11:

0449 M: lui parlava italiano n' /non/ sapeva parlare inglese allora
0449 dice
0449 " scusi signore " dice " me /mi/ dica di chi 'e quella giacca "
0449 quello non capiva XXX dice * "what a you want ?" =

Exemple 12:

0044 A: * allora quando entri la vasca è di lungo cosi
0045 G: a destra
0045 A: * but all the way to the end =
0046 = * completamente in fondo

Exemple 13:

0064 A: the other two bedrooms are nice * perché una è
0065 G: and plus there is another bedroom [downstairs]
0065 A: * [tredici per tre-]

Exemple 14:

0955 S: je sais pas * ci dicem'a /diciamo/a Ghilma ma ma ma ma
0957 * enfin j'ai e.

L'ensemble de ces types d'alternance intra-phrastique touche 266 occurrences ou 5,5% du corpus. Cette catégorie se situe mi-chemin entre l'alternance entre deux phrases et celle entre des éléments faisant partie d'une même phrase. Enfin, la forme pure d'alternance intra-phrastique impliquant l'alternance entre syntagmes (ex. 15) ou éléments syntagmatiques (ex. 16) concerne 67 occurrences, soit 1,4% du corpus.

Exemple 15:

0160 R: * vostra mamma =
0161 = * dit que c'est pas bon

Exemple 16:

0476 S: la Grande Allée pis la rue St-Jean
0476 L: ben la Grande Allée ça c'est * the chic part * XXX =
0477 = * ça c'est la rue principale là j'pense

Bien que cette catégorie semble marginale par son nombre, elle demeure néanmoins la catégorie qui pourrait impliquer la plus grande maîtrise des langues concernées. Au total, les différents types de l'alternance intra-phrastique affectent donc 333 occurrences ou 6,9% du corpus.

Un dernier niveau d'analyse concerne l'emploi de mots isolés. L'ensemble des mots doit être analysé selon qu'il s'agit d'emprunts intégrés, d'emprunts momentanés ou d'alternances. Signalons qu'il s'agit d'une question d'un intérêt particulier car chacune de ces catégories peut avoir une signification différente par rapport à l'intégration sociolinguistique du locuteur. On y retrouve l'emploi de substantifs (ex. 17) regroupant 326 occurrences ou 6,7% du corpus, et l'emploi de mots de diverses catégories grammaticales (ex. 18) regroupant 97 occurrences ou 2,0% du corpus.

Exemple 17:

0338 A: c'est qui Serge ?
0339 D: ()
0339 A: ton * boyfriend *

Exemple 18:

0186 R: il aurait fallu aller chez monter à sept heures après le
0186 souper restait toute la journée ici avec les enfants a pouvait
0186 pas monter =
0187 = * allora * ça pas servi beaucoup le le le

Par ailleurs, à cette catégorie s'ajoutent 62 occurrences d'emplois para-linguistiques de mots de diverses catégories (ex. 19). En totalité, les unités lexicales comptent 485 occurrences dans le corpus, soit 10,0% de celui-ci.

Exemple 19:

4111 P: * coopérative * vuole dire

Enfin, l'emploi de noms propres utilisés dans une autre langue que celle de laquelle ils originent (ex. 20) totalise 475 occurrences.

Exemple 20:

1646 L: j'sais pas The Wai- Witness si c'est bon avec e.
1646 F: e. ça d'l'air bon avec * Harrison Ford *
1647 L: ouais mais y est pas aussi j'trouve cool quand y est toute
1647 clean cut lui

Il s'agit de 9,8% de la totalité des occurrences du corpus. En soi, cette catégorie semble peu pertinente puisqu'il faut bien nommer les gens ou les choses par leur nom. Les résultats laissent toutefois entrevoir quelques disparités si l'on tient compte des langues impliquées. Nous y reviendrons.

La variation en fonction des langues

Si l'on retient certaines des catégories décrites plus haut pour en analyser la répartition en fonction des langues, on verra que la proportion d'emploi des langues représente une autre dimension de la variation. Prenons, par exemple, l'emploi des noms propres:

itl < fr	67	
fr < itl	39	
itl < ang	90	
ang < itl	29	
fr < ang	219	46,1%
ang < fr	31	

On note à partir de ces résultats que les différences entre les langues peuvent être assez importantes. Dans ce cas-ci, l'utilisation de noms propres anglais lorsque le locuteur s'exprime en français représente près de la moitié des occurrences, soit 46,1%. Ceci peut être attribuable à l'univers culturel dans lequel vivent les locuteurs puisque c'est souvent en parlant de sociétés, de produits de consommation, de cinéma ou de sports que l'on cite des noms en anglais.

Un autre exemple est celui de l'emploi de mots:

itl < fr	121	24,9%	
fr < itl	44		
itl < ang	77		
ang < itl	17		--------
fr < ang	214	44,1%	
ang < fr	12		----- 6,0%

Le phénomène le plus fréquent est l'utilisation de mots anglais dans des énoncés français, soit près de la moitié des occurrences (44,1%), ce qui est probablement imputable en partie à la variété même de français utilisée à Montréal, variété qui comprend des recours à l'anglais et qui laisse la possibilité au locuteur d'avoir recours à des emprunts momentanés. Un autre phénomène notable est l'utilisation de mots français dans des énoncés italiens, un phénomène qui compte pour 24,9% de l'emploi de mots. Ceci pourrait être surtout attribuable à l'univers institutionnel et culturel dans lesquels les locuteurs vivent (des quartiers francophones, un État francophone). Enfin, un autre fait

est à noter: au total, seulement 29 énoncés anglais comprennent un mot exogène, qu'il soit italien ou français, ce qui correspond à 6,0% des intégrations lexicales. Il est vrai que l'anglais est la langue qui a été la moins utilisée en général mais même en procédant à une pondération des résultats[4], la proportion d'emploi de mots italiens ou français en anglais demeure très inférieure. Il semblerait que dans l'usage qu'en font les italophones de Montréal, l'anglais soit moins perméable à l'intégration lexicale que les autres langues.

Un dernier exemple illustrant la variation en fonction des langues est celui de l'alternance intra-phrastique pure: ce type d'alternance met en cause des éléments liés syntaxiquement (syntagmes ou éléments syntagmatiques):

itl < fr	22		
fr < itl	15		
		37	55,2%
itl < ang	7		
ang < itl	4		
		11	16,4%
fr < ang	12		
ang < fr	7		
		19	28,4%

On note que les alternances intra-phrastiques pures, c'est-à-dire à l'intérieur d'une phrase entre des syntagmes ou des éléments syntagmatiques, sont produites dans plus de la moitié des cas lorsque le français et l'italien sont impliqués ensemble, ce qui devra être étudié plus à fond. On peut émettre l'hypothèse que la proximité de ces deux langues jouerait un rôle dans cette situation. Ainsi, la contrainte d'équivalence syntaxique pourrait être plus facilement appliquée en fonction de ces deux langues.

La variation en fonction des locuteurs

Ces deux dimensions des choix linguistiques des italophones de Montréal que sont les catégories conversationnelle, linguistique ou lexicale d'une part, et les langues dans lesquelles elles prennent forme, d'autre part, révèleront tout leur sens lorsqu'elles seront analysées en fonction des caractéristiques sociales des locuteurs qui en font usage. Il s'agit là de l'étape subséquente à l'établissement des faits bruts et c'est précisément la direction dans laquelle cette recherche doit être poursuivie. Pour tenter de comprendre comment s'articulent les choix linguistiques en situation de plurilinguisme, c'est sûrement la sociolinguistique qui a su développer les instruments les plus adéquats au cours des dernières années. Un problème de taille subsiste cependant lorsque la variabilité concerne autant d'options. Dans la recherche de la compréhension globale des comportements linguistiques en société, il faudra continuer à développer des instruments d'analyse devant faciliter l'établissement de rapports complexes entre les faits linguistiques et les faits sociaux.

---

[4]La pondération devrait s'effectuer comme suit. Les locuteurs ont produit leurs tours de parole dans des proportions de 50,4% en italien, 29,8% en français et 19,8% en anglais. Leur utilisation de mots exogènes dans un tour de parole italien est de 198 occurrences. Elle est de 258 occurrences dans un tour de parole français et de 29 occurrences dans un tour de parole anglais. En ramenant la proportion des tours de parole à 33% pour chaque langue, on obtiendrait 129,6 mots exogènes dans l'italien (28,0%), 285,7 mots exogènes dans le français (61,6%) et 48,3 mots exogènes dans l'anglais (10,4%).

RÉFÉRENCES BIBLIOGRAPHIQUES

AUER, P. (1984), *Bilingual conversation*, Philadelphie, John Benjamin.

GUMPERZ, J.J. (1982), "Conversational code switching", dans *Discourse strategies*, Cambridge, Cambridge University Press, 55-99.

LABRIE, N. (1987), "Les comportements langagiers des italophones de Montréal: un modèle d'analyse linguistique", dans *Revue québécoise de linguistique théorique et appliquée*, 6:2, 9-23.

POPLACK, S. (1980), "Sometimes I'll start a sentence in English y termino en español: Towards a typology of code-switching", dans *Linguistics*, 18, 581-618.

POPLACK, S. et SANKOFF, D. (sous presse), "Code-switching", dans U. Ammon, N. Dittmar et K.J. Mattheier (Eds), *Sociolinguistics. An international handbook of the science of language and society*, Berlin, New York, Walter de Gruyter Verlag.

POPLACK, S., SANKOFF, D. et MILLER, C. (1987), *The social correlates and linguistic consequences of lexical borrowing and assimilation*, Montréal, Centre de recherches mathématiques, Technical Report n° 1454.

SANKOFF, G. (1980), "Language use in multilingual societies: some alternate approach", dans *The social life of language*, Philadelphie, University of Pennsylvania Press, 29-46.

TABLEAU

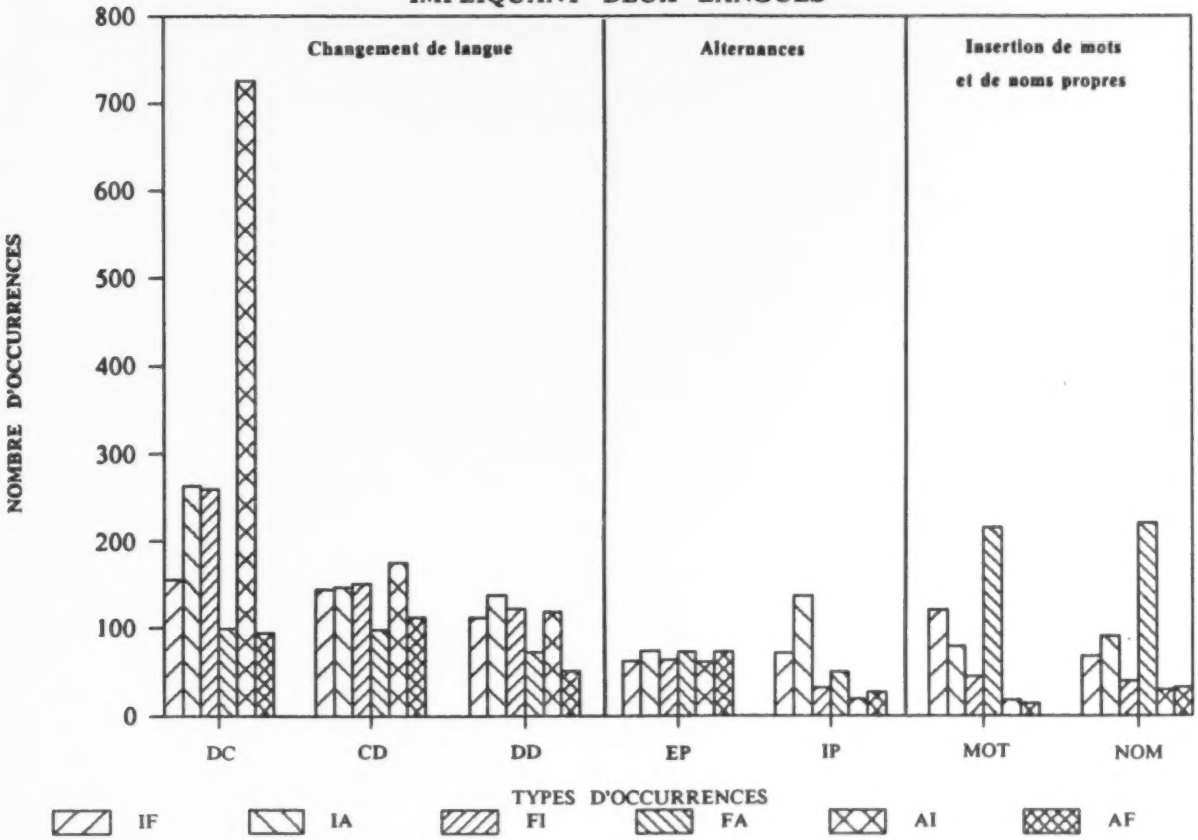
LE CORPUS DES ITALOPHONES DE MONTRÉAL
TABLEAU SYNOPTIQUE DES MANIFESTATIONS DU CONTACT DES LANGUES

Type	Nombre	S-TOTAL	%
Conversationnel			
1. DC (diverge/converge)	1596		32,9%
2. CD (converge/diverge)	824		17,0%
3. DD (diverge/diverge)	731		15,1%
		3151	65,0%
Extra-phrastique			
1. 2 phrases - nouvel interlocuteur	79		1,6%
2. 2 phrases - même interlocuteur	283		5,8%
3. discours rapporté "p1"	16		0,3%
4. faux départ	26		0,5%
		404	8,3%
Noms propres			
1. nom propre		475	9,8%
Mots			
1. diverses fonctions:			
1. substantif	326	326	6,7%
2. verbe	26		
3. adjectif	23		
4. conjonction	14		
5. préposition	8		
6. nominalisateur	6		
7. ponctuant	11		
8. adverbe	2		
9. déterminant	7		
		97	2,0%
2. paralinguistique	62		1,3%
		485	10,0%

Intra-phrastique					
	1. interjection au début		**94**		
	2. syntagme libre		**81**		
	3. discours rapporté p1:"p2"		**32**		
	4. 2 prop. coord.	**1. conjonction La**	**3**		
		2. conjonction Lb	**5**		
		3. p1(interj.+ conj.)*p2	**5**		
	5. 2 prop. subord.	**1. nominalisateur La**	**4**		
		2. nominalisateur Lb	**17**		
	6. apposition	**1. traduction+répétition**	**15**		
		2. non reprise	**7**		
		3. expression/explicitation	**3**		
				266	**5,5%**
	7. analyses				
	P				
		1. SN * SV	5		
		2. SP * [P] SN	6		
		3. SV * SP	10		
		4. SV * SN	2		
		5. SP * SV	1		
		6. SP * SP	1		
		7. SN * SP	2		
		8. SN * SN	1		
	SV				
		9. V * SP	2		
		10. Vcop * SP	2		
		11. V * [P] SN	7		
		12. Vcop [adv] * SN	10		
		13. V * adj	1		
		14. aux * part	1		
	SN				
		15. dét * [dét] N	5		
		16. N * Mod	1		
		17. SN * Mod	5		
	SP				
		18. prép*[P]/[prép dét]N	2		
		19. prép * SN	2		
	SA				
		20. adj * SP	1		
				67	1,4%
				333	6,9%
				4848	

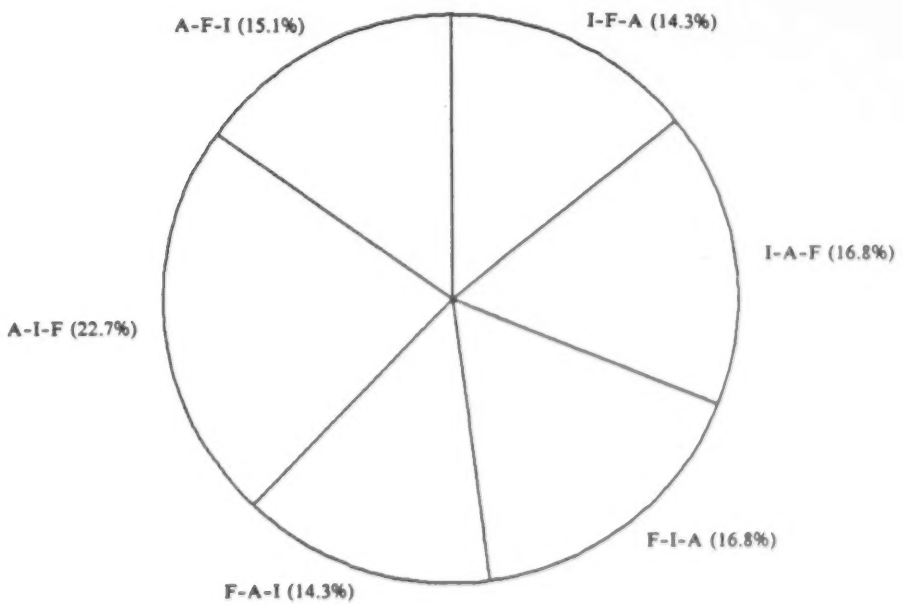
DISTRIBUTION DES OCCURRENCES

IMPLIQUANT DEUX LANGUES



DISTRIBUTION DES 119 OCCURRENCES

TRILINGUES DE TYPES CONVERSATIONNEL (DD)



Pi INTERROGATIF EN QUÉBÉCOIS

Gilberte Léger
Université de Sherbrooke

0. Présentation

Il existe, en québécois comme dans d'autres parlers régionaux, de petits mots qui fusent à l'oral, mais qui apparaissent rarement dans la grammaire scolaire, dans les dictionnaires. Qui d'entre nous n'a jamais entendu d'*entécas*, de *fak*, de *coudon*, de *voyons don*, de *tiens* et, bien sûr, de *pi* en conversant avec un locuteur québécois! Tous ces petits mots, je les nommerai, en empruntant le terme à O. Ducrot, les "mots du discours".

Ils sont intéressants et méritent d'être étudiés car, d'une part, ils font partie de l'idiolecte d'un bon nombre de locuteurs québécois et donc s'inscrivent dans le parler régional en question. D'autre part, peu d'études ont cerné la valeur et le rôle de ces unités en contexte, ce qui ouvre les portes à tout un courant d'études et de réflexion portant sur le langage authentique.

Bien que la langue soit considérée comme un fait oral par les linguistes, les grammaires sont généralement axées vers une description de la langue écrite (ou plus ou moins orale), prescrivant et proscrivant l'usage. Ainsi ont été façonnés les systèmes d'après cet objet déjà normé: les faits de l'oralité qui ne correspondent pas à la norme écrite établie sont souvent rejetés ou marginalisés, peut-être faute d'outils aptes à en rendre compte. Par exemple, les Québécois ont appris à "traduire" *pi* par *et*, *puis*, etc. à l'écrit. Mais ce qui est fascinant, c'est que malgré tout, pi, comme beaucoup d'autres unités typiques de l'oral, demeure usité encore aujourd'hui en québécois.

1. Problématique

Pi versus pi?

Pi est une unité suffisamment complexe pour justifier plusieurs études (celles de P. Laurendeau 1982, 1983 le démontrent), et les énoncés 1, 2, 3, et 4, tirés d'un de ses articles, illustrent différents rôles de *pi*:

1. Pierre pi Jean mangent.
   [Addition de notions (substitution par et)]

2. I a parlé aux élèves pi aux professeurs; ou plutôt aux professeurs en premier.
   [Succession temporelle (substitution par puis)]

3. J't'aime pi j't'en veux.
   [Opposition (substitution par mais)]

4. I est en esti pi i sourit (pareil).
   [Opposition (substitution par mais)]

Dans ces exemples comme dans bien d'autres, *pi* peut obtenir un statut théorique et s'intégrer, de fait, dans la grammaire: c'est alors une conjonction de coordination dont la valeur se précise en contexte linguistique. On donne à *pi* le statut de connecteur, car il est utilisé pour permettre une "connexion", un lien entre des catégories.

Mais bon nombre d'occurrences de *pi* observées en discours spontané ne peuvent être considérées ici. Et c'est en examinant *pi* qu'il a été possible d'isoler les emplois non conformes de *pi*?: les cas qui m'intéressent ne correspondent pas à la structure *X pi X*[1], mais plutôt à *PD*$_1$ *pi? PD*$_2$[2].

Pi n'a donc pas pour unique fonction de lier des catégories. Il peut tout aussi bien relier des portions de discours

5. PD$_1$ thème = une grève.
   - Pi? (Qu'est-ce qui s'est passé?)
   PD$_2$ - Ben là, les flics sont arrivés, les gars ont frappé pi...

qu'amorcer un échange ou encore signifier le désir du locuteur de clore le dialogue (*-pi? PD*$_1$ */ PD*$_1$ *pi*?):

6. (Salut!) Pi? Comment ça va?
   Qu'est-ce que tu fais de bon?

7. - J'ai vu Martin hier soir.
   - Pi? / Pi après? (Qu'est-ce que tu veux que ça m'fasse?)

Dans le premier cas il est connecteur discursif (désormais CD), car il intervient au niveau thématique en marquant une relation sémantique entre *PD*$_1$ et *PD*$_2$ (voir exemple 5). Il est aussi connecteur discursif lorsque l'énoncé avec *pi*? ouvre le dialogue sur un thème précis:

8. Pi(?) tes examens?

Dans le deuxième cas, *pi*? est marqueur discursif (MD) et ne vise aucun lien thématique entre des portions de discours. Les locuteurs l'utilisent soit pour amorcer une conversation en laissant à l'interlocuteur le choix du thème,

9. (Salut!) Pi?

soit pour signifier un reproche, un refus, un manque d'intérêt, etc.

---

1 X représente une quelconque catégorie. J'entends par "catégorie" les constituants des syntagmes (N., ADJ., ADV., V.), les constituants phrastiques (SN, SV...), ainsi que les propositions entières (P) qui se manifestent sous forme d'énoncés.

2 PD$_1$ et PD$_2$ correspondent à différentes portions de discours.

10. - Ils viennent de pogner le petit dernier de la voisine à piquer au dépanneur du coin.
- Pi après? / Pi? / Pi ça? (C'est pas de nos affaires. / Qu'est-ce que ça peut ben faire?)

Les différentes approches s'élaborent à partir de leur objet d'étude. Certains phénomènes langagiers caractéristiques de l'oral (les mots du discours, par exemple) n'ont été que peu étudiés jusqu'à récemment, et je retiens des paramètres qui me permettront, je l'espère, de cerner le rôle, le fonctionnement et la valeur de *pi*? en contexte. J'ai donc considéré, à l'intérieur des limites de la linguistique, les paramètres suivants: la syntaxe et la sémantique. Et comme je travaille sur une unité typique du québécois oral, je crois indispensable d'élargir ces paramètres au discours. Il sera donc question de syntaxe discursive et de sémantique discursive.

2. Étude de *pi*?

2.1 Syntaxe

Il s'agit de considérer une syntaxe distributionnelle (sans base catégorielle / fonctionnelle) qui permette de percevoir les règles d'emploi, soit de situer physiquement *pi*? dans le discours. Les locuteurs québécois utilisent *pi*? dans des conditions, avec des valeurs et à des fins précises, et je crois qu'il existe une syntaxe caractérisante des emplois de *pi*? qui servira mes objectifs de description.

Je suggère les paramètres syntaxiques suivants:

- La position de l'énoncé avec *pi*? dans le discours (en 2.1.1)
- La compatibilité entre *pi*? et les catégories (en 2.1.2)
- Les combinaisons avec *pi*? (en 2.1.3)

2.1.1 La position de l'énoncé avec *pi*? dans le discours

Ici, 3 positions retiennent l'attention:

- l'attaque de discours, soit le début de l'échange.
- l'interne de discours: *pi*? se situe dans le corps de l'échange.
- la conclusion de discours.

En attaque de discours, *pi*? ouvre le dialogue. On l'utilise simplement pour amorcer le discours, sans thème précis (voir exemple 6), ou encore pour lier l'énoncé qu'il introduit à des connaissances communes aux interlocuteurs. Ainsi un thème est proposé, dès le début de l'échange (voir exemple 8).

À l'interne de discours, *pi*? est utilisé en réaction à une portion de discours qui précède. Il peut s'agir d'une interrogation stricte (voir exemple 5), ou d'une interrogation pouvant être considérée comme rhétorique (voir exemple 7). Nous verrons en 2.2.1 l'implication de cette distinction.

Ce qui caractérise la conclusion de discours, c'est qu'à première vue aucun emploi de *pi*? n'y est possible. Effectivement, *pi*? CD (demande de suite) n'est jamais utilisé dans le but de conclure l'échange. Et le *pi*? MD (marquant un reproche, un refus, un désintérêt), laisse à l'interlocuteur le pouvoir de justifier ses propos, d'argumenter, etc...

2.1.2 Le deuxième paramètre syntaxique retenu permet de voir la compatibilité entre *pi*? et les catégories. Le relevé des catégories compatibles avec *pi*?, soit l'entourage catégoriel immédiat, contribue aussi à la description de *pi*?.

Il est utile d'envisager ce paramètre autant des points de vue de l'autonomie de **pi?** (voir exemples 5 et 9), que de la compatibilité (et parfois même de la dépendance (pi X?)) entre **pi?** et une catégorie (voir exemples 6 et 11 respectivement).

11. PD_1 thème: Marthe et ses projets de voyage
Pi Charles? (Qu'est-ce qu'il va faire?)

En attaque de discours, *pi*? peut apparaître seul (*-pi*?) ou être accompagné d'une catégorie (*-pi X*?).

À l'interne de discours, *pi*? peut aussi être autonome et constituer l'énoncé en termes de formes linguistiques. Toutefois, certaines contraintes apparaissent selon sa valeur, son rôle discursif: il arrive que *pi*? ne puisse apparaître seul, comme c'est le cas lors d'un changement ou d'un glissement de thème. L'énoncé a alors obligatoirement la structure *pi X*?, où X est le nouveau thème (ou propos) introduit (voir exemple 11).

Il est aussi intéressant de constater que *pi*? MD et les syntagmes figés *pi après*? / *pi ça*? sont toujours autonomes. Le locuteur peut, bien sûr, faire suivre son énoncé *pi après*? / *pi ça*? ou *pi*? d'un autre énoncé, mais il y aura alors une pause délimitant les énoncés (voir exemple 10).

2.1.3 Le 3^{e} paramètre syntaxique retenu touche les combinaisons avec *pi*?. J'ai identifié 2 types de combinaisons avec *pi*?. Les combinaisons du 1er type se composent en associant librement *pi*? et une (ou plusieurs) autre(s) unité(s). Elles répondent au besoin de précision qu'ont parfois les locuteurs, permettant d'éviter une ambiguïté potentielle .

12. Sur la liste, il y a du savon, des tomates, des piments...
- Pi ensuite (de ça)?
- Pi après (ça)?

Mais ce sont les combinaisons du 2^{e} type qui m'intéressent vraiment. Ce sont des syntagmes désormais figés, qui correspondent à des unités, tant syntaxiquement que sémantiquement (les unités regroupées sont indissociables sur le plan syntaxique; ces syntagmes ont un sens précis). Ce sont *pi après*? et *pi ça*? (voir exemple 10). Ces derniers ont, comme particularité syntaxique, de ne jamais apparaître en attaque de discours. Et à l'interne, ils sont toujours autonomes.

Déjà, certaines distinctions apparaissent. Mais la similitude distributionnelle observée à plusieurs reprises limite la caractérisation (par opposition) des emplois et des valeurs contextuelles de *pi*?.

La syntaxe permet donc de faire des classes, de cerner certaines propriétés de *pi*?.

2.2 Sémantique

Je propose maintenant une analyse sémantique où la valeur interrogative de l'énoncé (stricte ou rhétorique), de même que la notion de thème, sont des paramètres qui permettront de saisir la valeur de *pi*? en contexte.

2.2.1 Interrogation stricte et interrogation rhétorique

Faire une interrogation stricte, c'est poser une véritable question. Et en posant la question, le locuteur attend généralement une réponse à la question de la part de l'interlocuteur.

13. - J'ai vu le dernier film de Woody Allen hier soir.
- Pi? (c'était-tu bon?)

Mais fréquemment, à l'interne de discours, les énoncés avec *pi*? sont en fait des interrogations rhétoriques. Par "interrogation rhétorique", j'entends un énoncé qui, de par sa forme linguistique, constitue une interrogation (marquée par un point d'interrogation à l'écrit, par l'intonation à l'oral, par la structure syntaxique de l'énoncé...), mais où le locuteur n'attend pas de l'interlocuteur qu'il réponde directement à la question posée. Il vise autre chose, comme par exemple reprocher, signifier un refus, marquer un désintérêt (voir exemples 7 et 10).

J'illustre tout d'abord ce fait par l'exemple bien connu "Peux-tu me passer le sel?" ou encore, si je dis "Il fait chaud dans cette salle. Est-ce que quelqu'un peut ouvrir une fenêtre?". L'interprétation stricte de ces interrogations porterait l'interlocuteur à répondre "oui", "non", "peut-être ben". Mais c'est habituellement l'interprétation rhétorique de ce genre d'énoncés qui est retenue: on nous passera le sel ou quelqu'un ira ouvrir la fenêtre.

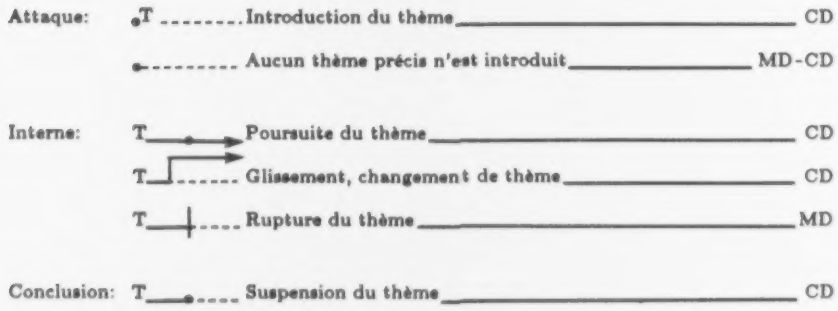
Si on applique maintenant le concept d'interrogation rhétorique aux énoncés avec *pi*?, à l'interne de discours, il est évident que le locuteur qui énonce *pi après*? / *pi ça*? ne demande aucune information supplémentaire. Il veut plutôt signifier à son interlocuteur qu'il n'est nullement intéressé par le thème proposé, par la tournure de la conversation ou il marque son désaccord. Et sans entrer dans les détails, je dirai simplement que les causes du manque d'intérêt du locuteur peuvent être de plusieurs ordres, comme par exemple le refus volontaire d'engager la conversation suite à une querelle; le reproche d'aborder tel thème, de telle façon, etc.

L'intérêt de distinguer l'interrogation stricte et l'interrogation rhétorique est donc de permettre d'expliquer sémantiquement le comportement syntaxique de l'unité.

2.2.2 Notion de thème

La notion de thème est centrale, nécessaire dans la détermination des valeurs attribuables à *pi*? en contexte discursif. Elle concerne l'orientation du discours et permet, à l'interne surtout, de séparer nettement les emplois de *pi*? CD et de *pi*? MD, ce que j'illustre comme suit:

PI? ET L'ORIENTATION THÉMATIQUE DU DISCOURS



Le *pi*? CD s'associe à l'interrogation stricte et le *pi*? MD à l'interrogation rhétorique.

À l'attaque, *pi*? peut tout aussi bien annoncer, introduire un thème spécifique, qu'indiquer à son interlocuteur qu'il lui laisse le soin d'ancrer la conversation sur un thème de son choix. Et bien qu'il paraisse étrange, à première vue, qu'un énoncé où *pi*? est autonome soit considéré "ancreur thématique", plusieurs exemples attestent son rôle sémantique et le rendent plausible: le thème est si évident que l'amorce avec *pi*? y réfère implicitement.

14. (Madeleine sait que Jacques revient d'une compétition de natation. Elle amorcera l'échange comme suit:)
- Pi? (ta compétion)

À l'interne, *pi*? CD marque un lien thématique et favorise un développement du thème, ou rend acceptable un changement de thème (voir exemples 5 et 11). Il constitue une vraie demande axée sur un thème précis[3]. Dans le 1er cas, on réfère au thème antérieur:

15. - PD_1 (T_1)
- Pi? (Qu'est-ce qui s'est passé?) T_1
- PD_2 (T_1)

et dans le 2e cas, on se réfère au nouveau thème introduit:

16. - PD_1 (T_1)
- Pi Marie? Qu'est-ce qu'elle devient? T_2
- PD_2 (T_2)

C'est d'ailleurs la notion de thème qui rend le mieux compte de la contrainte syntaxique -*pi X*? (X obligatoire) lors d'un énoncé introduisant un changement de thème.

---

3 T_1 = premier thème introduit.
T_2 = deuxième thème introduit.

À l'interne toujours, *pi*? MD a pour rôle d'indiquer que le locuteur refuse une élaboration du thème ou les propos de l'interlocuteur. Ainsi, l'énonciateur de *pi*? MD conteste la pertinence du thème ou des propos (voir exemples 7 et 10).

17. - PD_1 (T_1)
    - Pi? (Qu'est-ce que tu veux que ça me fasse?) T_1
    (- PD_2)

Conclusion

L'étude détaillée de *pi*?, dont j'ai présenté ici un aperçu, permettra de démontrer l'intérêt et l'importance d'intégrer des paramètres de diverses natures dans l'étude des faits oraux. La syntaxe seule ne permet pas de caractériser *pi*?, de rendre compte de sa valeur dans ses emplois de CD et de MD. La sémantique seule ne suffit pas non plus à distinguer certains emplois de *pi*? CD à l'attaque et à l'interne de discours. Allier syntaxe et sémantique me semble donc le moyen le plus adéquat de caractériser *pi*?.

D'autres paramètres, également pertinents dans l'étude de *pi*? mais délaissés dans ce texte pour plusieurs raisons (contraintes de temps, d'espace...), seront utilisés ultérieurement. La prise en compte des actes illocutoires stricts et dérivés, des effets perlocutoires visés ou atteints par exemple n'est pas à négliger.

BIBLIOGRAPHIE

Ducrot, O. et al. (1980), *Les mots du discours*, Paris, Les Éditions de minuit.

Laurendeau, P. (1982), "Pi I et Pi 2 en français du Québec: un continuum sémantique à désambiguiser", dans *Revue de l'Association québécoise de linguistique*, vol. 2, pp. 91-108.

Laurendeau, P. (1983), "Sur la systématique et la combinatoire du joncteur pi en québécois", dans *Travaux de linguistique québécoise*, Collection Langue française au Québec, section 4, Les presses de l'Université Laval.

Léard, J.M. (1984), "Contraintes sur la subordination et la coordination en dialogue: de la syntaxe à la sémantique", communication présentée à l'ACFAS, Québec, Université Laval, à paraître en version modifiée dans *Langue française*, sous le titre: "Dialogue et connecteurs propositionnels: syntaxe, sémantique et pragmatique".

Lyons, J. (1980), *Sémantique linguistique*, Larousse, Collection Langue et langage.

Searle, J.R. (1972), *Les actes de langage*, Paris, Collection Savoir, Hermann.

ANALYSE PARAMÉTRIQUE DE LA CHUTE DU /S/ EN ESPAGNOL[1]

Valérie Catrice
Université du Québec à Montréal

0. Introduction

Le but de cette recherche est de clarifier le rôle de la structure métrique et prosodique sur la chute du /s/ final de mot dans quatre dialectes espagnols américains, soit ceux de la République Dominicaine, de la Bolivie, du Vénézuela et de Cuba. Nous étudierons plus particulièrement l'influence de facteurs linguistiques tels que l'accentuation, le contexte segmental suivant et la longueur du mot sur le phénomène d'affaiblissement du /s/.

On trouve le /s/ final de syllabe en position interne et en position finale de mot:

(1) position interne: escribir "écrire"

position finale: mes "mois"
casas "maisons"

Le /s/ peut être lexical comme dans MES "mois" en (1); il peut aussi avoir une fonction grammaticale. Il représente le suffixe pluriel des déterminants, noms et adjectifs dans le syntagme nominal:

(2) las casas rojas "les maisons rouges"

Il représente aussi le suffixe de la deuxième personne du singulier dans le syntagme verbal:

(3) cantas "(tu) chantes"

Certains auteurs tiennent compte de la fonction grammaticale du /s/ dans leur explication du phénomène d'amuïssement de ce segment. Ils proposent, dans le cas du /s/ pluriel, une hypothèse fonctionnelle définie en termes de conservation d'information en structure de surface: l'information de pluralité pertinente y sera retenue (cf. Ma et Herasimchuk 1971, Cedergren 1973, Terrell 1975 et 1977, Poplack 1979, Alba 1982a). Alba (1982b) introduit la notion d'accentuation du mot: les mots appartenant à une catégorie majeure (substantifs et adjectifs) sont accentués et les mots appartenant à une catégorie mineure (déterminants), atones. Ces derniers favorisent la conservation du /s/ tandis que les mots toniques ont tendance à perdre leur /s/.

---

[1]Cet article est un résumé du mémoire de maîtrise présenté par l'auteure à l'Université du Québec à Montréal.

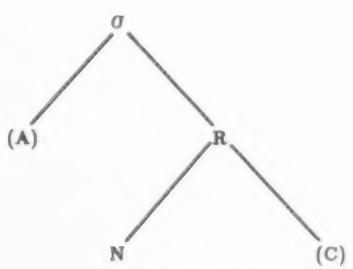
Malgré une similarité des effets de conditionnement dans plusieurs dialectes espagnols, révélée par des études quantitatives, il existe quelques différences mineures d'une région dialectale à une autre quant à l'influence possible des facteurs phonologiques (Terrell, 1978).

1. Cadre théorique

Dans cette section, nous traiterons de la notion d'accentuation au niveau de la syllabe car nous pensons que l'accentuation syllabique a une influence directe sur l'élision du /s/ final de mot. Nous adopterons une approche métrique, où l'accentuation est définie en termes de proéminence relative entre les syllabes (Liberman et Prince, 1977). D'après ces auteurs, les syllabes, les mots et les syntagmes sont organisés en une structure rythmique hiérarchique.

La syllabe, définie comme étant un pic de sonance compris entre des segments d'un degré de sonance moins élevé (Harris, notes de cours)[2], a une structure interne. Celle-ci joue un rôle important car les règles phonologiques, comme le placement de l'accent par exemple, sont sensibles à l'organisation interne des constituants de la syllabe. Nous supposons la structure syllabique suivante (cf. Harris, 1983):

(4)



où σ = syllabe
( ) = constituant optionnel
A = attaque
R = rime
N = noyau
C = coda

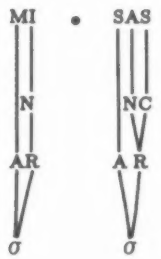
La syllabe se divise en deux constituants immédiats, l'attaque (constituant optionnel) et la rime (constituant obligatoire). On retrouve dans l'attaque tous les segments précédant le pic de proéminence syllabique (le noyau). La rime, elle, comprend tous les autres segments.

Le seul élément obligatoire est la rime, qui comprend un constituant obligatoire, le noyau, et un constituant optionnel, la coda.

La structure interne d'un mot comme MISAS "messes", par exemple, sera:

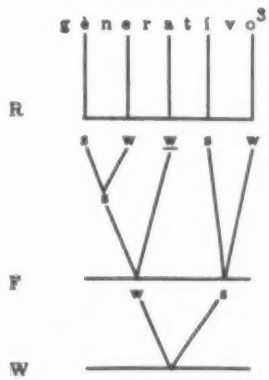
[2]Cf. Catrice (1987) pour une discussion de la définition de la syllabe, section 1.2.

(5)



Les règles d'assignation de l'accentuation sont sensibles à la structure interne de la rime. Il faudra donc spécifier que les règles d'accentuation ne s'appliquent que sur la projection de la rime ("rime projection", Hayes (1985)), une représentation ne comprenant que les rimes de la chaîne phonologique. Les rimes seront regroupées de droite à gauche en des constructions binaires, les pieds métriques, étiquetées S (w) et sujettes à la condition de branchement ("Branching Condition"): "Footnodes labeled w(eak) cannot branch." (Harris, 1983: 111 (5.44)). Les pieds simples restants seront adjoints, au niveau du mot, comme membre faible du pied binaire adjacent déjà formé (cf. (6), w). Au niveau du mot, les pieds seront regroupés en des arbres branchant à droite, étiquetés (w) s.

(6)



(Harris, 1983: 96 (5.8)

où: R = niveau de la rime
F = niveau du pied
W = niveau du mot
\ = accent non primaire
/ = accent primaire
w = faible
s = fort

L'accent primaire tombera sur la syllabe qui n'est dominée que par des noeuds étiquetés s (la syllabe *tí* en (6)).

[3]Note de l'éditrice: les accents indiqués par l'auteure correspondent à des accents phonétiques et il peut arriver que ceux-ci ne correspondent pas aux acccents graphiques de l'espagnol: dans ce cas, par exemple, l'orthographe d'usage du mot GENERATIVO ne comprend aucun accent.

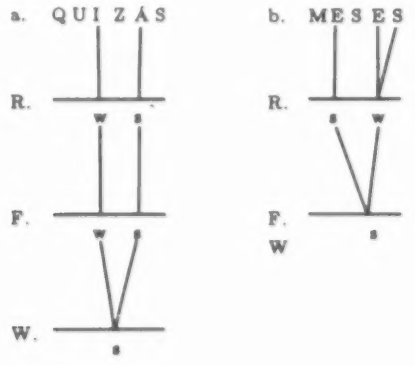
Les mots de fonction atones, c'est-à-dire les déterminants (LOS "les", MIS "mes"), les prépositions (A "à", DE "de"), les pronoms (QUE "que", LOS "ceux"), les clitiques objets (LOS / LAS/ LES "les"), etc., se cliticiseront sur le mot tonique suivant pour former ainsi une unité prosodique accentuée, car tout mot atone peut se cliticiser (Zwicky, 1977).

1.1 Prédictions

Dans cette section, nous appliquerons ce modèle d'accentuation au phénomène d'élision du /s/ final de mot.

On retrouve le /s/ final de mot tant dans une syllabe tonique (QUIZAS "peut-être") que dans une syllabe atone (MESES "mois (pluriel)"). Leur représentation métrique sera:

(7)



En (7a.), le /s/ final de mot se trouve dans un pied simple fort (niveau F) et en (7b.), le /s/ final est dans la branche la plus faible d'un pied fort (niveau F).

Nous aimerions suggérer qu'une branche faible d'un pied fort à l'extrême droite du mot perdra son /s/ plus fréquemment que si cette branche est l'unique branche d'un pied fort (donc étiquetée s): le /s/ final de mot chutera donc plus souvent dans un mot comme MESES "mois (pluriel)" que dans un mot comme QUIZÁS "peut-être". Il semblerait donc que l'élision du /s/ pluriel, généralement dans une syllabe atone, sera plus fréquente que celle d'un /s/ lexical, généralement dans une syllabe tonique. Un /s/ lexical se trouvant dans une syllabe atone chutera aussi souvent que le /s/ pluriel dans le même contexte. C'est le cas du /s/ final du mot ENTONCES "alors", par exemple[4].

Nous aimerions aussi suggérer que le contexte dans lequel apparaît le /s/ influence la chute de ce segment. Nous proposons donc un modèle à trois paramètres linguistiques, soit:

[4]Cf. Catrice (1987) section 1.1 pour une discussion des prédictions faites pour les clitiques objets et les déterminants.

(8) segmental: contexte segmental suivant

---Consonne
---Voyelle
---Pause

accentuel: contexte suivant et précédant la syllabe porteuse du /s/;

syllabe porteuse du mot

longueur du mot: mot porteur du mot

D'après ce modèle, l'élision du /s/ résultera de l'interaction de ces trois paramètres linguistiques, tant dans le syntagme nominal que dans le syntagme verbal.

Il semble y avoir une corrélation entre la position du /s/ final de mot et sa fonction grammaticale (c'est-à-dire /s/ lexical, pluriel et verbal): son appartenance à une syllabe atone ou tonique dans un environnement donné favorisera la rétention ou l'élision de ce segment.

Quant à la variation dialectale, nous aimerions suggérer qu'elle se situe à deux niveaux: celui de l'organisation interne des facteurs considérés et celui du taux de chute du /s/ final de mot.

2. Méthodologie

2.1 Corpus

Les données de cette recherche proviennent de quatre enregistrements de langage spontané d'environ une demi-heure, de locuteurs hispanophones représentant quatre régions différentes de l'Amérique du Sud et des Caraïbes, soit la République Dominicaine, la Bolivie, le Vénézuela et Cuba. Une transcription de chaque entrevue a été faite pour pouvoir ensuite relever tous les cas de /s/ en position finale de mot. Ceux-ci, tant lexicaux que grammaticaux, font partie des deux types de syntagmes (nominal et verbal).

Dans le syntagme nominal, les /s/ lexicaux et pluriels font partie des noms (MES "mois", CASAS "maisons"), des adjectifs (DIEZ "dix", TODOS "tous", AZULES "bleus"), des déterminants (UNOS "quelques", LOS/LAS "les", MIS "mes"), des pronoms (NOSOTROS "nous", LOS "ceux"), et du quantifieur MAS "plus". Dans le syntagme verbal, les /s/ font partie des verbes (ESTÁS, "(tu) es"), des auxiliaires (HAS IDO "(tu) as été), des pronoms objets clitiques (LOS TOMAMOS "(nous) les prenons") et des adverbes (DESPUES "après").

Certains cas de /s/ final de mot n'ont pu être retenus lors de la codification des données pour diverses raisons: le contexte précédant ou suivant le mot porteur du /s/ n'a pu être relevé; le mot porteur du /s/ était au début d'une phrase non terminée. Par exemple, dans la phrase: Entonces, y despues vamos... "Alors, et après (nous) allons...", le /s/ de l'adverbe ENTONCES ne sera pas retenu. Les expressions figées telles que DIGAMOS "disons", TU VES QUE "tu vois que", etc., n'ont pas été relevées.

Un /s/ en position finale de mot a plusieurs variantes (ou plusieurs représentations phonétiques), comme par exemple la sibilante [s], la variante aspirée [h] et l'élision complète Ø. Dans le cadre de cette recherche, nous nous sommes surtout intéressée au contexte d'élision du segment. La seule distinction entre les variantes que nous avons faite a été entre la présence du segment sous l'une ou l'autre des variantes et l'absence complète du /s/.

Les taux de chute du /s/ final pour les quatre locuteurs sont représentés au tableau 1:

TABLEAU 1:

Taux d'élision du /s/ dans les quatre dialectes espagnols américains dans les deux types de syntagmes (nominal et verbal)

dialectes syntagmes	République Dominicaine	Bolivie	Vénézuela	Cuba
syntagme nominal	70% (271)	60% (284)	67% (207)	57% (91)
	N=386	N=474	N=311	N=160
syntagme verbal	61% (112)	42% (80)	49% (81)	60% (80)
	N=183	N=192	N=167	N=133

N = Nombre total de données dans le syntagme

2.2 Groupes de facteurs

Parmi les facteurs linguistiques étudiés, nous retrouvons l'accentuation de la syllabe (syllabe porteuse du /s/, syllabe suivante et syllabe précédente) et celle du mot (mot atone VS mot accentué), le contexte segmental suivant et la longueur du mot porteur du /s/. Nous ne considérons ici qu'un seul facteur extra-linguistique, soit celui de la région dialectale.

Nous avons classé les /s/ selon leur appartenance à l'un ou l'autre des deux syntagmes (nominal ou verbal) et non pas selon leur fonction grammaticale (lexical, pluriel ou verbal).

Nous aurons donc la classification suivante:

(10) CODIFICATION GÉNÉRALE POUR LES DEUX SYNTAGMES

Groupes de facteurs: 6

1. Contexte segmental suivant[5]

- consonne	las casas	"les maisons"
- voyelle	las amigas	"les amies"
- pause	casas #	"maisons"

2. Accentuation de la syllabe porteuse du /s/

- atone	lăs mansanăs	"les pommes"
- tonique	estás	"(tu) es"
	˘ = atone	/= tonique

3. Accentuation de la syllabe suivante

- atone	las bŏtellas	"les bouteilles"
- tonique	los párques	"les parcs"
- pause	estas #	"(tu) es"

4. Accentuation de la syllabe précédente

- atone	logicas	"logiques"
- tonique	papéles	"papiers"
- pause	# los	"les"

5. Catégorie syntaxique du mot porteur du /s/

- mot de fonction atone	los, mis	"les", "mes"
- mot de fonction tonique	unos, esos	"quelques", "ces"
- mot plein tonique	estamos	"(nous) sommes"
	casas	"maisons"

6. Nombre de syllabes dans le mot porteur du /s/

- mot monosyllabique	dos	"deux"
- mot polysyllabique	casas	"maisons"

2.3 Analyse statistique

Nous avons effectué une analyse statistique sur les données pour étudier l'influence de l'interaction des groupes de facteurs entre eux sur l'amuïssement du /s/. Un algorithme (VARBRUL 2S) calculera la signification des paramètres en (10) à l'aide d'une analyse multivariée: un indice de vraisemblance maximale calculé pour chacun des groupes de facteurs sera comparé à l'indice de vraisemblance maximale original, calculé pour toutes les données. Les groupes de facteurs significatifs seront choisis par ordre d'importance (cf. Cedergren, 1983; Poplack, 1979).

[5]Les contextes suivants et précédents se réfèrent aux contextes suivant et précédant la syllabe porteuse du /s/.

Dans un premier temps, nous avons traité les données des quatre dialectes séparément. Puis nous avons fait une analyse conjointe de toutes les données, en spécifiant à quel dialecte appartenaient les données. Une analyse statistique a été refaite sur ces données pour étudier les tendances générales.

3. Résultats: analyse et discussion

Nous présenterons dans cette section les résultats des différentes analyses statistiques. Après avoir regardé l'évaluation des paramètres pour chacun des dialectes, nous tenterons de clarifier les limites de la variation dialectale à l'aide d'une comparaison des résultats des quatre dialectes. Nous regarderons d'abord, pour chaque cas, les résultats du syntagme nominal puis ceux du syntagme verbal.

3.1 Syntagme nominal

Le tableau 2 comprend les résultats du syntagme nominal pour les quatre dialectes. Il est à noter que dans tous les tableaux, les valeurs plus petites que .5 ($p < .5$) favorisent l'élision du /s/ et les valeurs plus grandes que .5 ($p > .5$), la rétention du segment. Les valeurs égales à .5 représentent des tendances neutres.

3.1.1 Résultats pour chaque dialecte

(i) République Dominicaine

L'accentuation de la syllabe porteuse du /s/, la catégorie syntaxique du mot porteur du /s/ et l'accentuation de la syllabe suivante sont les trois seuls groupes de facteurs significatifs quant à la chute du /s/: une syllabe porteuse du /s/ atone (.061), les mots accentués (pleins (.245) et de fonction (.428)) de même qu'une syllabe suivante atone (.354) favorisent cette élision du /s/ final de mot.

(ii) Bolivie

Pour ce dialecte, les groupes de facteurs significatifs sont l'accentuation de la syllabe porteuse du /s/, le contexte segmental suivant et la longueur du mot porteur du /s/. Une syllabe atone favorisera la chute du /s/ (.157) de même qu'un mot polysyllabique (.389). Quant au contexte segmental suivant, il n'y a que la pause qui favorise la rétention du /s/ (.754).

(iii) Vénézuela

Trois groupes de facteurs ont été choisis comme étant significatifs par VARBRUL 2S, soit le nombre de syllabes dans le mot porteur du /s/, l'accentuation de la syllabe suivante et l'accentuation de la syllabe porteuse du /s/. Un mot polysyllabique aura tendance à perdre son /s/ (.331); il en est de même pour une syllabe porteuse du /s/ atone (.341). Quant à l'accentuation de la syllabe suivante, une syllabe atone et une pause favorisent l'élision du segment.

TABLEAU 2:

Groupes de facteurs significatifs influençant la chute du /s/ dans le syntagme nominal de quatre dialectes espagnols américains

Dialectes Groupes de facteurs	République Dominicaine	Bolivie	Vénézuela	Cuba
Accentuation de la syllabe porteuse du /s/	at. .061[1] ton. .939	at. .157[1] ton. .843	at. .341[3] ton. .659	
Catégorie syntaxique du mot porteur du /s/	pac .245[2] fac .428 fat .804			
Accentuation de la syllabe suivante	at. .354[3] ton. .506 pa. .640		at. .388[2] pa. .418 ton. .687	
Nombre de syllabes dans le mot porteur du /s/		poly .389[3] mono .611	poly .331[1] mono .696	
Contexte segmental suivant		-V .321[2] -C .408 -P .754		-C .313[1] -V .482 -P .703
Accentuation de la syllabe précédente				ton. .329[2] at. .671
Signifiance	p < .035	p < .001	p < .001	p < .001
N	396	474	311	160

pac = mot plein acentué; fac = mot de fonction accentué; fat = mot de fonction tonique; poly = mot polysyllabique; mono = mot monosyllabique; -V = voyelle; -C = consonne; -P = pause; 1,2,3 = ordre de sélection des groupes de facteurs significatifs; at. = atone; ton. = tonique; pa. = pause; N = nombre de données.

(iv) Cuba

Ce quatrième dialecte se démarque des trois autres par son choix des groupes de facteurs significatifs. En effet, les deux seuls groupes de facteurs retenus sont le contexte segmental suivant et l'accentuation de la syllabe précédente. Comme pour le dialecte du Vénézuela, l'effet du contexte segmental suivant semble se situer au niveau de la présence ou de l'absence d'une syllabe suivante: le /s/ chutera plus souvent s'il est suivi d'une syllabe débutant par une consonne (.313) ou une voyelle (.482) que s'il est suivi d'une pause (.703). Quant à l'accentuation de la syllabe précédente, une syllabe tonique favorise l'élision du /s/ (.329) et une syllabe atone, la conservation de ce segment (.671).

3.1.2 Comparaison dialectale

Parmi les groupes de facteurs significatifs (tableau 2), quatre paraissent plus importants. Il s'agit de l'accentuation de la syllabe porteuse du /s/, nettement plus significative, de la longueur du mot porteur du /s/, de l'accentuation de la syllabe suivante et du contexte segmental suivant.

Il semble y avoir un effet de compensation entre le contexte segmental suivant et l'accentuation de la syllabe suivante: lorsque le premier est choisi (Bolivie et Cuba), le second est rejeté. Cet effet de compensation semble vouloir suggérer que la syllabe suivante exerce une influence sur la chute du /s/ final de mot, soit par son accentuation, soit par le contexte segmental.

Les deux derniers groupes de facteurs choisis, soit la catégorie syntaxique du mot porteur du /s/ et l'accentuation de la syllabe précédente, ont une importance moindre: ils ne sont retenus que pour un seul dialecte chacun.

3.1.3 Tendances générales

Une deuxième analyse statistique a été effectuée sur l'ensemble des données combinées. On pourra alors étudier l'effet général des groupes de facteurs sur la chute du /s/ final de mot.

TABLEAU 3:

Groupes de facteurs significatifs influençant la chute du /s/ dans le syntagme nominal de quatre dialectes espagnols américains. Combinaison des données.

accentuation de la syllabe porteuse du /s/	nombre de syllabes dans le mot	accentuation de la syllabe suivante	distinction dialectale	accentuation de la syllabe précédente
at. .257	poly .408	at. .368	Vén. .389	ton. .384
ton. .743	mono .592	ton. .504	Répd. .412	at. .517
		pa. .628	Bol. .566	pa. .600
			Cuba .632	
signifiance p < .045	N = 1331			

at. = atone; ton. = tonique; pa. = pause; poly = mot polysyllabique; mono = mot monosyllabique; Vén. = Vénézuela; Répd. = République Dominicaine; Bol. = Bolivie; N = nombre de données.

Deux groupes de facteurs n'ont pas été choisis comme étant significatifs, soit le contexte segmental suivant et la catégorie syntaxique du mot porteur du /s/.

Le groupe de facteurs le plus significatif quant à la chute du /s/ semble bien être l'accentuation de la syllabe porteuse du /s/: un /s/ chutera plus fréquemment dans une syllabe atone (.257) que dans une syllabe tonique (.743).

Un mot polysyllabique, ainsi qu'une syllabe suivante atone et une syllabe tonique, favorisent l'élision du /s/.

D'après les valeurs de la distinction dialectale, les dialectes du Vénézuela et de la République Dominicaine favorisent l'élision du /s/ (.389 et .412 respectivement) et les deux autres dialectes (Bolivie et Cuba), la conservation du segment (.566 et .632 respectivement). Il semble y avoir une concordance entre cet ordre et celui obtenu au tableau 1, soit:

République Dominicaine	>	Vénézuela	>	Bolivie	>	Cuba
70%		67%		60%		57%

Il semble en effet que la République Dominicaine et le Vénézuela favorisent la chute un peu plus que les deux autres dialectes.

L'influence de la distinction dialectale semble résider aussi au niveau du choix des groupes de facteurs significatifs: les trois premiers groupes de facteurs choisis sont ceux retenus pour le dialecte du Vénézuela (tableau 2), dialecte favorisant le plus la chute du /s/ (tableau 3).

L'importance du facteur accentuel au niveau syllabique dans le syntagme nominal se réflète dans le choix des groupes de facteurs significatifs: l'accentuation syllabique dans tous les contextes (syllabe porteuse du /s/, contexte précédant et suivant celle-ci) a été retenue.

3.2 Syntagme verbal

Les résultats du syntagme verbal se trouvent au tableau 4:

TABLEAU 4:

Groupes de facteurs significatifs influençant la chute du /s/ dans le syntagme verbal de quatre dialectes espagnols américains

dialecte	République Dominicaine	Bolivie	Vénézuela	Cuba
groupes de facteurs				
accentuation de la syllabe porteuse du /s/	at. .100 to. .900	at. .262 to. .738	at. .320 to. .680	at. .328 to. .672
signifiance	p < .001	p < .001	p < .001	p < .001
N	183	192	167	133

N = nombre de données; at. = atone; to. = tonique

3.2.1 Résultats pour chaque dialecte

Un seul groupe de facteurs est retenu comme étant significatif pour les quatre dialectes, soit l'accentuation de la syllabe porteuse du /s/: pour chacun des dialectes, une syllabe atone favorise l'élision du segment.

3.2.2 Comparaison dialectale

Les résultats du tableau IV montrent l'importance du facteur accentuel au niveau de la syllabe sur la chute du /s/ pour les quatre dialectes. Non seulement l'accentuation de la syllabe porteuse du /s/ est l'unique groupe de facteurs retenu, mais elle est aussi très significative dans les quatre dialectes ($p < .001$).

3.2.3 Tendances générales

Nous avons effectué une seconde analyse statistique sur l'ensemble des données combinées du syntagme verbal. Les résultats de cette analyse se trouvent au tableau 5:

TABLEAU 5:

Groupes de facteurs significatifs influençant la chute du /s/ dans le syntagme verbal de quatre dialectes espagnols américains. Combinaison des données.

accentuation de la syllabe porteuse du /s/		contexte segmental suivant		distinction dialectale	
atone	.241	---V	.348	Cuba	.407
tonique	.759	---C	.454	Répd.	.458
		---P	.693	Vén.	.530
				Bol.	.605
signifiance = p < .003			N = 675		

---V = voyelle; ---C = consonne; ---P = pause; Répd. = République Dominicaine; Vén. = Vénézuela; Bol. = Bolivie; N = nombre total des données du syntagme verbal.

Seulement trois groupes de facteurs ont été retenus comme étant significatifs quant à la chute du /s/, soit l'accentuation de la syllabe porteuse du /s/, le contexte segmental suivant et la distinction dialectale.

La chute du /s/ dans une syllabe atone sera favorisée (.241). Quant au contexte segmental suivant, il semble que ce soit la présence d'une syllabe suivante, indépendamment du contexte segmental suivant, qui favorise l'élision du segment. En effet, une voyelle et une consonne favoriseront la chute du /s/ et la pause, la rétention du segment. Il existe néanmoins une différence dans le taux de chute selon le type de segment (---V (.348) > ---C (.454)).

Le /s/ chute plus fréquemment dans les dialectes de Cuba et de la République Dominicaine (.407 et .458 respectivement) que dans les deux autres dialectes (Vénézuela .530 et Bolivie .605). Nous avons un ordre similaire au tableau I quant au taux de chute du /s/ dans le syntagme verbal des quatre dialectes:

République Dominicaine	>	Cuba	>	Vénézuela	>	Bolivie
61%		60%		49%		42%

3.3 Comparaison des deux syntagmes

Nous comparerons, dans cette section, le choix des groupes de facteurs significatifs pour les deux syntagmes pour ensuite essayer d'en tirer quelques conclusions.

Le facteur accentuel au niveau de la syllabe porteuse du /s/ semble jouer un rôle important dans le phénomène d'élision. C'est le seul facteur significatif dans le syntagme verbal des quatre dialectes, et le plus significatif dans le syntagme nominal: il est retenu pour trois dialectes. Une syllabe atone favorisera dans tous les cas la chute du /s/ final. Dans le syntagme nominal, d'autres groupes de facteurs influencent l'élision du segment, soit l'accentuation de la syllabe suivante, le contexte segmental suivant et la longueur du mot porteur du /s/. L'accentuation de la syllabe précédente et la catégorie syntaxique du mot porteur du /s/ ne sont significatifs que pour un dialecte chacun.

Nous remarquons, dans le tableau 6 (p. 97), que pour le syntagme verbal, il semble qu'un seul facteur soit significatif dans la chute du /s/ tandis que dans le syntagme nominal, plusieurs facteurs interagissent entre eux pour influencer l'affaiblissement du segment.

4. Conclusion

L'élision du /s/ final de mot semble être conditionnée par l'environnement: nous avons pu constater l'effet des facteurs suggérés (c'est-à-dire les facteurs accentuel, segmental et la longueur du mot porteur du /s/) sur la chute du segment dans les deux syntagmes.

Le facteur accentuel semble être le plus significatif, surtout au niveau de la syllabe porteuse du /s/: la condition pour que le /s/ soit élidé est l'absence d'accentuation de cette syllabe.

La classification des données a été faite selon leur appartenance à un syntagme particulier et non selon leur fonction grammaticale. Il semble que ce soit bien l'accentuation de la syllabe porteuse du /s/ qui régisse l'élision ou la rétention du segment et non la fonction grammaticale de celui-ci: un /s/ pluriel chutera plus souvent qu'un /s/ lexical car ce dernier se trouve généralement dans une syllabe tonique et le /s/ pluriel, dans une syllabe atone:

(11)	/s/ pluriel:	MESES	"mois (pluriel)"
		LOS/LAS	"les"
		ESTOS	"ces"
	/s/ lexical:	DOS	"deux"
		QUIZÁS	"peut-être"
		MES	"mois"

Le /s/ verbal chutera s'il est dans une syllabe atone (TENEMOS "(nous) avons") tandis qu'il sera conservé plus volontiers dans une syllabe tonique (ESTÁS "(tu) es", ES "(il/elle) est").

Il existe néanmoins quelques variations dialectales qui se situent à deux niveaux. Tout d'abord, dans l'organisation interne des groupes de facteurs conditionnant l'amuïssement du /s/ et dans le choix des groupes de facteurs significatifs quant à la chute du segment. Nous avons vu qu'il semble y avoir plus de variation dans le syntagme nominal que dans le syntagme verbal (cf. tableau VI). Enfin, les facteurs n'ont pas toujours la même influence dans tous les dialectes.

La variation dialectale se situe aussi au niveau du taux de chute du /s/: dans deux dialectes (République Dominicaine et Bolivie), le /s/ aura la même tendance dans les deux syntagmes (nominal et verbal). Dans le dialecte de la République Dominicaine, le /s/ aura tendance à être élidé dans les deux syntagmes et dans le dialecte de la Bolivie, le segment sera plus souvent conservé dans les deux syntagmes. Nous retrouverons, dans les deux autres dialectes une différence dans les deux syntagmes quant aux tendances de chute du /s/.

En général, dans un même dialecte, l'élision du /s/ sera plus favorisée dans le syntagme nominal que dans le syntagme verbal, sauf dans le dialecte de Cuba qui montre l'effet contraire (cf. tableau 1).

TABLEAU 6:

Combinaison des groupes de facteurs significatifs influençant la chute du /s/ dans les deux syntagmes (nominal et verbal) de quatre dialectes espagnols américains

dialectes / groupes de facteurs	République Dominicaine	Bolivie	Vénézuela	Cuba
Accentuation de la syllabe porteuse du /s/	X Y	X Y	X Y	 Y
Catégorie syntaxique du mot porteur du /s/	X			
Accentuation de la syllabe suivante	X		X	
Contexte segmental suivant		X		X
Nombre de syllabes dans le mot porteur du /s/		X	X	
Accentuation de la syllabe précédente				X

X = Groupes de facteurs significatifs pour le syntagme nominal
Y = Groupes de facteurs significatifs pour le syntagme verbal

BIBLIOGRAPHIE

ALBA, O. (1982a), *Estratificación social del español de Santiago de los Caballeros: La S implosiva*, Thèse de doctorat inédite, Université de Puerto Rico.

ALBA, O. (1982b), "Función del acento en el proceso de elisión de la S en la República Dominicana", dans O. Alba (ed), *El español del Caribe: ponencias del VI simposio de dialectología*, Universidad Católica Madre y Maestra, pp. 15-26.

CATRICE, V. (1987), *Analyse paramétrique de la chute du /s/ en espagnol*, Mémoire de maîtrise inédit, Université du Québec à Montréal, 61 pages.

CEDERGREN, H.J. (1973), "On the nature of variable constraints", dans C.J. Bailey et R. Shuy (eds), *New ways of analyzing variation*, Georgetown University Press, Washington D.C., pp. 13-22.

CEDERGREN, H.J. (1983), "Sociolingüística", dans H. López Morales (coordinator), *Introducción a la lingüística actual*, Editorial Playa, pp. 147-165.

HARRIS, J.W. (1983), *Syllable structure and stress in Spanish: a non-linear analysis*, Linguistic Inquiry monograph 8, Cambridge, MIT Press, 158 pages.

HARRIS, J.W. (1987), *Notes de cours*, Université de Barcelone, Barcelone.

HAYES, B. (1985), *A metrical theory of stress rules*, Thèse de doctorat, MIT, Bloomington, Indiana University Linguistics Club, 212 pages.

LIBERMAN, M. et PRINCE, A. (1977), "On stress and linguistic rhythm", dans *Linguistic Inquiry*, vol. 8, n° 2, pp. 249-336.

MA, R. et HERASIMCHUK, E. (1971), "The linguistic dimensions of a bilingual neighborhood", dans Fishman, Coper, Ma (eds), *Bilingualism in the Barrio*, The Hague, Mouton, pp. 347-464.

POPLACK, S. (1979), *Function and process in a variable phonology*, Thèse de doctorat, University of Pennsylvania, 271 pages.

TERRELL, T.D. (1975), "Functional constraints on deletion of word final S in Cuban Spanish", dans *Berkely Linguistic Society*, vol. I, Berkely, University of California, pp. 431-437.

TERRELL, T.D. (1977), "Aspiration and deletion of word final /s/ in the Spanish of Caracas, Venezuela: a preliminary report", manuscrit, University of California, Irvine.

TERRELL, T.D. (1978), "Hacia un modelo de comparación dialectal: aspiración y elisión de S", manuscrit, University of California, Irvine.

ZWICKY, A. (1977), *On clitics*, reproduced by the Indiana Linguistic Club, 40 pages.

LEXICAL VARIATION IN THE SPEECH OF FOUR CHINESE SPEAKERS

Xu Daming
Université d'Ottawa

A well-accepted notion of linguistic variation is that "there are many alternate, semantically equivalent ways of saying 'the same thing'" (Labov, 1982: 18). Therefore, a (socio-) linguistic variable consists of variant forms which are identical in reference or truth value but opposed in their social and / or stylistic significance (Labov, 1972: 271). However, this definition of (socio-) linguistic variable has proven to be more successful with phonological studies than with the non-phonological ones. As Lavandera (1978) points out, there is a series of methodological difficulties in defining non-phonological variables with the semantic requirement. Some of Lavandera's (1978) arguments can be re-stated as follows:

As the non-phonological linguistic categories (such as words, syntactic constructions, etc.) tend to have referential meanings which are not so readily accepted as identical, the identification of a variation context where the forms under study mean exactly the same thing is frequently a rather arbitrary decision which lacks intersubjective agreements. Moreover, the identification of such a context necessarily excludes from the analysis large sets of data which also involve the forms in question. The result is that the scope of the study and the significance of the study are both restricted.

To the above points I would like to add that the semantic equivalence requirement of linguistic variables (as well as being a defining feature of linguistic variation) unnecessarily excludes meaning from the scope of variation study. As meaning is invariably held constant as a controlled variable, it has never been used as an explanatory factor whose weight in conditioning variation is studied.

To solve the problems she had raised, Lavandera (1978) suggests that the semantic equivalence requirement should be relaxed. Besides, in discussing the concept of syntactic variation, D. Sankoff holds the view that "distinctions in referential value or grammatical function among different surface forms can be neutralized in discourse" (Sankoff, 1986: 17). Therefore, his notion of linguistic variation involves "functionally" as well as referentially equivalent structures. However, to my knowledge, the above-mentioned ideas have not been executed in many studies of language variation. Most studies reported are still confined to the (controversially or uncontroversially) "semantically equivalent" cases. The discoursal neutralisation of functional or referential values of lexical items is still yet to be illustrated.

The research I will be reporting in the following is an experiment to illustrate such a neutralisation and one to dispense with the semantic restriction in an investigation of sociolinguistic variation. In a study of lexical composition of samples of natural speech, I defined a context for lexical variation in a semantically-undiscriminative environment at the discoursal level (a stretch of speech 250 word-tokens long). With such a definition of the range of variation (250 slots for any number of tokens of any type), the actual occurrences of various lexical items are taken as dependent variables in a quantitative analysis (what types in what token number filling the 250 slots). The results of the analysis show that the usage of words correlates with a number of linguistic and socio-demographic factors. A tentative conclusion, therefore, is that it is possible and sometimes necessary to define a

linguistic variable in a purely formal manner in order to avoid difficult or unnecessary semantic discrimination. Without the semantic restriction, the overall configuration of a speech corpus becomes of relevant interest in research, instead of only certain selected cases from the corpus. This does not only result in an economic utility of data, it also reduces arbitrariness and increases objectivity in the research methodology. Moreover, a formally-defined linguistic variable gives more flexibility in the study of the conditioning effects of various linguistic and sociological factors in that it allows meaning to be one of the varying interactive factors of the analysis.

Methodology

The speech data analyzed in the study are taken from tape-recorded conversations of four adult Chinese speakers in their natural interactive situations. All the speakers are professionals and they all speak the standard dialect (Putonghua). However, two of them were originally from Beijing; the other two were originally from the southern dialect areas. Sex is equally matched for the dialect-background groups (one male and one female for each group).

Two samples of speech are taken from each speaker, each from a different interview. Thus eight samples of speech are transcribed altogether. All samples are taken so that they begin at a turn-take (when the speaker started to speak during the conversation) and are truncated at the end of the 250th word-token (inclusive) of a stretch of continuous, uninterrupted speech. These restrictions on the samples are based on the following point and fact: 1) the spontaneity of the speech is emphasized in this study -- speech produced under real-time pressure is supposedly more consistently reflective of speakers' lexical proficiency, which falls within the scope of the study; the exclusive use of continuous and uninterrupted speech is meant to minimize any prompting or reflection factors on lexical generation; 2) in the speech corpus of recorded-interviews of continuous and uninterrupted speech, it is hard to find such a stretch longer than 250 word-tokens in all the eight needed recordings.

To study the effect of meaning difference on lexical variation, that is, to what degree the actual lexical composition of a stretch of speech is determined by the referential content of that stretch of speech, I made up, for this particular study, the operational definitions: "topic", and "content" vs. "function" words.

"Topic" is a representation of the referential meaning of a speech sample. Therefore, each speech sample is distinct from the others in "topic". As a premise, content words are much more topic-bound than function words. Therefore, any parameters in content and function word-type and word-token configurations are indications of the strength of topic on lexical variation.

Since the determination of a Chinese lexical item as whether belonging to one (or more) of the three parts of speech, noun, verb, and adjective or not is basically uncontroversial even in a context-free manner[1], I take it as the criterion to code all my word-tokens as either a "content" word (belonging to the noun, verb and adjective group) or a "function" word (not belonging to the noun, verb, and adjective group). The real meaning of a "content" word thus coded is that it is unam-

[1]To my knowledge, any finer classification would be theoretically controversial and context-sensitive and subject to inter-speaker interpretive variation for a large number of cases.

biguously referential as in contrast to a "function" word in the same coding, which may be disputedly or undisputedly recognized as containing only or mostly functional meaning and little referential meaning.

With a FORTRAN programme specifically developed for this study[2], I was able to let the computer do much of the indexing, sorting, counting, calculating and permuting work which is necessary for the quantitative analysis. The mechanical work thus becomes accurate and less labour- and time-consuming. This made it possible to make a few input and output manipulations to experiment with research alternatives.

In processing the data, I did not only input the eight original samples (operationally-defined as "topic-samples"), I also inputed merged samples. Specifically, I merged two samples of a same speaker into a speaker sample, two speaker samples into a group sample, and the whole data set into a "total" sample respectively.

In this study of lexical variation, the concept of "lexical sharedness" was adopted from Poplack, Sankoff, & Miller (1987). In my study, lexical sharedness is the indication of the recurrence of a word type[3] (or word types) in two or more speech samples by the same speaker or by different speakers. To obtain parameters of "lexical sharedness" for the purpose of examining associations between social factors and lexical variation, I also inputed matched samples. Matched-sample input is virtually inputting two or more separate samples' word-type list (the output of separate sample processing)[4].

For each separate sample input I obtained a word-type list in the order of frequency rank (types having more tokens are ranked higher) and a corresponding percentile table of frequency distribution. The word-type / word-token ratios are also calculated. For each matched-sample input, I obtained the shared-type / total-type ratio and the corresponding word-type list. For more-than-two-sample matches, I also obtained the sharing indice for all the word-types on the output list and a corresponding percentile table indicating their distribution.

For studying the topic effect on lexical variation, I obtained all the above-mentioned results separately for content words, for function words, as well as for all words of each input sample.

Findings

1. Proportions of content and function words

Table 1 shows the statistics of content and function-tokens in all the separate sample inputs. It is apparent that content word-tokens consistently occupy a smaller proportion in the net word-produce of oral speech consisting of 250 or more word-tokens. It is so across topics, across speakers, and across groups. The differences between the means of content and function-tokens are all highly significant, while the variations across samples are non-significant.

---

[2]I want to thank Youwen Wu for developing the program and for her helpful comments in the subject of the research.

[3]In definition, all occurrences of a given word are considered tokens of the same lexical type.

[4]The virtual inputs of separate samples are tokens; while the virtual inputs of matched samples are types.

The above-mentioned results can be taken as the first symptom that the choice of words in natural oral speech is not overwhelmingly determined by the referential meaning of the speech ("topic", as defined in this study). As the results show, it is the less referential group of words ("function" words, as defined in this study) which figure a larger proportion in the total word-tokens produced during a continuous span of speaking.

TABLE 1

Word-tokens

INPUT SAMPLES	CONTENT WORDS	FUNCTION WORDS	TOTAL N
TOPIC-SAMPLES			
BeijingMale1	46% (116)	54% (134)	250
BeijingMale2	48% (121)	52% (129)	250
SouthernMale1	46% (118)	53% (132)	250
SouthernMale2	47% (117)	53% (133)	250
SouthernFemale1	50% (126)	50% (124)	250
SouthernFemale2	47% (118)	53% (132)	250
BeijingFemale1	43% (108)	57% (142)	250
BeijingFemale2	48% (120)	52% (130)	250
MEAN	47%	53%	$p < 0,01$
SPEAKERS-SAMPLES			
BeijingMale	47% (237)	53% (263)	500
SouthernMale	47% (235)	53% (265)	500
SouthernFemale	49% (244)	51% (256)	500
BeijingMale	46% (228)	54% (272)	500
MEAN	47%	53%	$p < 0,01$
GROUP-SAMPLES			
Southern	48% (481)	52% (519)	1000
Female	47% (472)	53% (528)	1000
Mix (BMSF)	47% (465)	54% (535)	1000
Beijing	47% (465)	54% (535)	1000
Mix (BFSM)	48% (481)	52% (519)	1000
Male	47% (474)	53% (526)	1000
MEAN	47%	53%	$p < 0,01$
TOTAL-SAMPLE	47% (944)	53% (1056)	2000 $p<0,05$

2. Word-type / word-token ratios

Table 2 (p. 104) gives all the percentages of total word-types over total word-tokens in separate input samples and their means at different merged-sample levels. The number of tokens for each cell corresponds exactly to Table 1, so it is not repeated. It can be noted:

1) As the samples grow larger, the percentages become smaller. This is fairly consistent throughout the table. It is true for all words jointly, and for content and function words separately.

2) For all the samples, the type / token ratio is higher for content words than for function words, the discrepancy is large (25%+) at all the four levels represented here: topic, speaker, group, and total.

The word-type / word-token ratios are the inverse ratios of repetition rates of word-types within a limited span of speaking. In either case, we also see from the ratios the relative diversity of word choice. For instance, a low word-type / word-token ratio indicates a high repetition rate of the words used (the high recurrence rate of the involved word-types) and it also indicates that the range of used word-types is relatively small.

Apart from 1) and 2) in the above, it can also be noted in the table that there is variation in the repetition rates among topics, among speakers, and among groups, and that there is variation in the relative ranges of chosen word-types in respect to content and function words at the three identified levels. These variations make up an intrinsic part of lexical variation, and would consist of the objects of an interesting study. However, I singled out 1) and 2) in the above to focus on the question of topic effects on lexical variation.

What 1) represents is that, as various topics are lumped together, the words used to talk about them do not become obviously more diversified (increase proportionately in the range of types). Instead, proportionately less word-types are used when more topics are included. This un-isomorphic relationship between lexical composition and speech content is by itself unremarkable, for it illustrates the well-accepted truth that, compared to all the meaning differences we can express verbally, the linguistic devices we can use to express them are rather limited. It is remarkable here in that it shows that the meaning conditioning on lexical variation is actually quantified at various levels. The quantitative results show that the diversity of word-types decreases from about 50% of word-tokens at the 250-token level as the token number increases[5]. The proportionate decrease in vocabulary range cannot be presumably explained by an increased similarity in topics, for different things do not become more similar when they are calculated together. The secret of course lies in the fact that the discrete semantic concepts are more likely to be repeated across topics, but that has already come into lexical semantics, which is qualitatively different from the referential meaning at the discourse level. The lexical variation studied here essentially involves how speakers dissect a particular "meaning" in real life into lexical concepts. This is "ways of speaking" in contrast to "things to say". The implication of 1) is that, as a formally-defined variable, word-type / word-token ratio gives us an opportunity to find that lexical-type range co-varies with topic-range only at a very low degree and in a small corpus of oral speech data. As soon as the corpus grows large, this co-variation is superseded by the disproportionate relationship between the limited vocabulary and the unlimitedly different things speakers wish to talk about.

What 2) represents is the fact that content words are less often repeated than function words. This also argues that topic-sensitivity is not the major character of lexical variation. For a topic-sample (it coincides in features with the concept of "idiolect" (Bloch, 1948): the utterances produced by one speaker at one time and possibly on one topic), it is the topic-neutral function words which

[5]The merged samples inadvertently amalgamate in the confounding factors of "prompting", "reflection", and individual variation which are purged at the topic-sample level for the concept of "lexical generation under real-time pressure". If these factors are controlled, I expect that the diversity rate would drop more dramatically as the number of tokens increases than is indicated in our table.

are repeated much more often than the topic-bound content words. Among speakers and among group of speakers, and within a speaker (across topics: speaking on different occasions and about different things, the lexical similarity thus found cannot be basically attributed to that "they have similar things to say" rather than that "they have similar ways of saying things".

TABLE 2

Word-type / Word-token ratios

INPUT SAMPLES	ALL WORDS	CONTENT WORDS	FUNCTION WORDS
TOPIC-SAMPLES			
BeijingMale1	47%	64%	33%
BeijingMale2	45	60	30
SouthernMale1	39	47	33
SouthernMale2	42	51	35
SouthernFemale1	50	64	36
SouthernFemale2	48	63	34
BeijingFemale1	45	64	31
BeijingFemale2	50	62	40
MEAN	45,8	59,4	34,0 p<0,05
SPEAKERS-SAMPLES			
BeijingMale	39%	57%	22%
SouthernMale	34	44	25
SouthernFemale	39	54	29
BeijingMale	39	55	26
MEAN	37,8	52,5	25,5 p<0,05
GROUP-SAMPLES			
Southern	31%	47%	17%
Female	32	48	18
Mix (BMSF)	33	51	16
Beijing	31	44	20
Mix (BFSM)	32	50	17
Male	30	43	18
MEAN	31,5	47,2	17,7 p<0,05
TOTAL-SAMPLE	26%	41%	13% p<0,05

3. Proportion and frequency correlation of word-types

Table 3 illustrates a negative correlation between proportions of types in a sample and numbers of tokens of a type. The numbers of tokens per type are actually frequencies of occurrence of a type in a sample. As can be seen, a small proportion of word-types are used in much higher frequency than the rest of word-types, the majority of word-types are used in very low frequency. This characteristic is more illustratively displayed by Figure 1, in which the proportion of types drops continuously with the ascending order of frequency but it drops rather steeply at the beginning and rather gently towards the end. This characteristic is basically the same across samples: practically all the cross-sample differences are statistically non-significant.

The above suggests the following linguistic universals:

1) in any continuous speech corpus of moderate length, word-types are not used in equal frequency;

2) frequencies are negatively correlated with proportions of word-types out of the used type total.[6]

TABLE 3

Proportion of types by frequency

INPUT SAMPLE	NUMBER OF TOKENS PER TYPE										TYPE TOTAL
	1+	2+	3+	4+	5+	6+	7+	8+	9+	10+	
Topic-samples											
BeijingMale1	42%	24%	17%	12%	6%	3%	2%			1%	118
BeijingMale2	37%	24%	16%	12%	7%	5%	4%	4%		2%	112
SouthernMale1	51%	34%	21%	13%	6%		5%	3%			98
SouthernMale2	49%	25%	19%	12%	9%	8%	6%	3%	2%		106
SouthernFemale1	34%	21%	16%	10%	8%	3%	2%	1%			126
SouthernFemale2	40%	24%	16%	10%	8%	3%	3%		1%		119
BeijingFemale1	42%	25%	14%	12%	6%	4%		4%	3%		113
BeijingFemale2	36%	20%	13%	8%	5%	3%			2%	2%	126
Mean	41.4	25.6	16.5	11.1	6.9	4.1	3.7	3.0	2.0	1.7	114.8
Speaker-samples											
BeijingMale	40%	25%	18%	12%	10%	9%	7%	6%	5%		193
SouthernMale	52%	33%	46%	17%	14%	12%	9%	6%	5%	3%	172
SouthernFemale	43%	30%	21%	16%	11%	8%	5%			4%	193
BeijingMale	42%	27%	18%	11%	8%	6%	5%	5%	4%	3%	197
Mean	44.3	28.8	20.8	14.0	10.1	8.8	6.5	5.7	4.7	3.3	188.8
Total-sample	45	31	25	18	14	13	11	9	7	7	525

[6] As a matter of interest, we processed two samples of speech from an English speaker. The results are that the frequency distribution characteristic is very similar and any differences in percentages of corresponding cells between those two samples and our Chinese data are found to be statistically non-significant (see Appendix II for the results of the English data. A serious postulation of such a linguistic universal, however, must wait for a broad-range cross-cultural study.

FIGURE 1

Type proportion and usage frequency correlation in the total-smaple (N = 525)

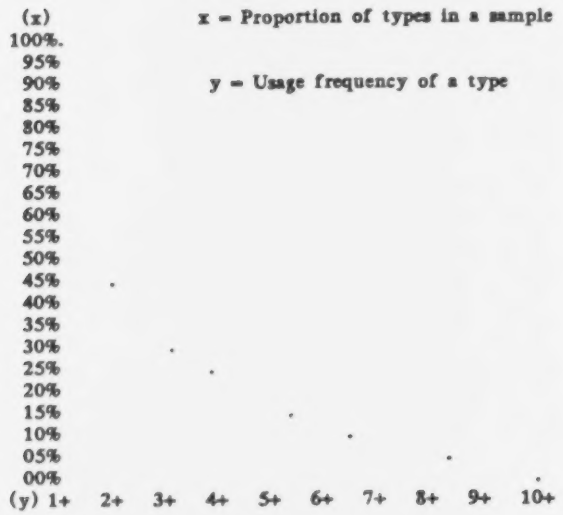


Table 4 gives similar results as Table 3 but for content words and function words separately. It can be seen that larger proportions of function words are used than content words in all the frequency distribution and the highest frequencies are exclusively occupied by function words. Thus we may add to the suggested linguistic universals that function words are repeated more often than content words and that the most frequently used words are function words. This has a particular implication for one of the questions under investigation: in the high frequency word-type variation, we can expect very little influence from topic conditions.

4. Sharing ratios

Table 5 (p. 109) gives the means of all the sharing ratios, which are part of the results of matched-sample inputs. The ratios are obtained by dividing the type totals of a separate sample with the shared-type total of a match including that sample. This way a ratio is obtained for each separate sample. As might be expected, since the sharing ratio of a sample is definitely influenced by the size of its matching sample(s), this influence would be a confounding factor to the ratios' indicativity of separate sample parameters. However, the factor is eliminated by taking the means of groups of ratios. For instance, the mean of all the ratios of inner-speaker topic-sample matches is taken to contrast with the mean of all the ratios of inter-speaker topic-sample matches. As both groups are matched in all possible combinations, the pairing-sample-size effects cancel each other when group means are taken.

TABLE 4

INPUT SAMPLE	NUMBER OF TOKENS PER TYPE										TYPE TOTAL
	1+	2+	3+	4+	5+	6+	7+	8+	9+	10+	
BeijingMale											
Content-word	31%	16%	10%	6%	4%	3%	1%	1%			136
Function-word	61%	46%	35%	28%	25%	23%	21%	18%	14%		57
SouthernMale											
Content-word	45%	24%	19%	12%	9%	8%	5%	3%	1%		105
Function-word	63%	46%	37%	25%	22%	18%	15%	10%		7%	67
SouthernFemale											
Content-word	32%	20%	11%	8%	4%	3%	2%				132
Function-word	67%	51%	41%	33%	28%	20%	13%			10%	61
BeijingFemale											
Content-word	35%	22%	13%	5%	2%	2%	1%				126
Function-word	56%	37%	27%	23%	18%	14%	13%	11%	10%	8%	71

The following can be noted from Table 5 (p. 109):

1) For two topic-sample matches, inner-speaker matches yield a significantly higher sharing ratio than the inter-speaker ones.

2) For the inter-speaker matches, the samples belonging to speakers from the same dialect group have a significantly higher ratio than those of different dialect groups.

3) For function words, topic samples have higher sharing ratios when the matches are among samples from speakers of same sex than from those of different sex.

4) For female speakers, matches among themselves yield higher ratios in content words and in all words than with samples from a male speaker, while the reverse is not true of male speakers when taken separately.

5) When topic-samples are matched in 8, the sharing ratios are much lower than when they are matched in 2.

6) Speaker-samples also have higher sharing ratios in same-dialect-match than in different-dialect-match.

7) Speaker-samples also have higher sharing ratios in same-sex-match than in different-sex-match.

8) Female speakers have significantly different ratios in inner-sex than in inter-sex match, while males do not have such significant differences (results thus not appearing in the table).

9) Two speaker matches have higher sharing ratios than the 4 speaker match.

10) When non-homogeneous groups are matched (Mix-Mix Match), a higher sharing ratio is yielded than the homogeneous group matches (Sex-group Match and Dialect-group Match).

11) Sex-group Match results in a higher sharing ratio than Dialect-group Match.

12) Throughout the table, sharing ratios for function words are consistently higher than those of the content words under the same conditions.

The interpretation of the above twelve points is as follows:

A. 1) shows that there is more lexical similarity between two stretches of speech from the same speaker than those from different speakers. This is to say that lexical variation is socially conditioned (individual variation).

B. 2), 3), 4), 6), 7), 8), and 10) show that lexical similarity also covaries with speakers' social attributes. Speakers share more word-types if they belong to a homogeneous social group, less if they do not. In 2), 3), 4), 6), 7), 8), this is indicated positively by the higher ratios of more homogeneous groups. In 10), the same trend is represented by the higher sharing ratio between non-homogeneous groups and by the lower sharing ratio between homogeneous groups respectively.

C. Relating the results noted in 2), 3), 4), 6), 7), 8) and 10) to each other, it can be inferred that dialect background is a more important factor in determining lexical sharedness. Also the sex factor is apparent in female lexical sharedness but is not for males.

D. 5) and 9) show that either the topics or the speakers, when taken together, do not have much in common. The similarity between either two topic-samples or between two speaker-samples is not the same similarity from match to match. However, even if the small proportion of lexical types shared by all may consist of a linguistic (syntactic, lexical, semantic) necessity, the remaining unshared proportion consists of the manifested lexical variation. This variation is much larger in proportion than the invariant part of natural-speech lexical composition. As all the preceding interpretation goes, this variation is not so readily explained by a meaning-form dependency than by a socially-conditioned commonality among speakers' active vocabularies.

E. 12) is the most meaningful for arguing against a topic-(dis)similarity explanation of lexical variation. The fact is that for any lexical similarity and dissimilarity among social groups, it is the function words which are more typically differentiating. The content words, supposedly more topic-sensitive, are always smaller in sharing ratios under the same condition and less differentiating between social groups. It is remarkable that the mean of content-type sharing ratios of the 8 topic-sample match is only 3%. That is an obvious indication that all the samples are different in topic, which in turn supports the claim in 2. that the proportionate decrease in lexical-type range as speech corpus grows larger is not because of increased similarity in topics.

TABLE 5

Means of sharing ratios[7]

INPUT SAMPLES	CONTENT WORDS	FUNCTION WORDS	ALL WORDS
2 TOPIC MATCH			
INNER-SPEAKER	22%	57%	35%
INTER-SPEAKER	18%	50%	28%
INTER-SPEAKER			
SAME DIALECT	20%	53%	32%
DIFF. DIALECT	17%	48%	29%
SAME SEX	N.S.	50%	N.S.
DIFF. SEX	N.S.	46%	N.S.
FOR FEMALE			
SAME SEX	21%	N.S.	34%
DIFF. SEX	16%	N.S.	30%
8 TOPIC MATCH	3%	18%	9%
2 SPEAKER MATCH			
SAME DIALECT	25%	57% (p<0,10)	36%
DIFF. DIALECT	20%	51% (p<0,10)	31%
SAME SEX	N.S.	54%	32% (p<0,10)
DIFF. SEX	N.S.	51%	31% (p<0,10)
FOR FEMALE			
SAME SEX	N.S.	53% (p<0,10)	34%
DIFF. SEX	N.S.	50% (p<0,10)	30%
4 SPEAKER MATCH	11%	36%	19%
GROUP MATCH			
Mix-Mix Match	27%	60%	36%
Sex-group Match	N.S.	N.S.	35% (p<0,10)
Dialect-grp Match	N.S.	N.S.	31% (p<0,10)

5. Shared types

If the occurrence of a lexical item is an absolute structural necessity, it must be included in the shared 18% of function words in the eight topic-sample match[8]. Therefore, the occurrences of at

[7]For all input samples, N > 203; the particular figures are already given in 2.2. p < 0,05 unless indicated otherwise.

[8]In fact, only 8 function types are shared in the match out of a total of 135 function types and a total of 525 types used in the total data (6% and 2% respectively).

least 82% of the function words in each topic-sample can be explained by social factors. One may argue that those occurrences should be best explained by a semantic necessity, that is, the meaning content of the speech determined which particular words are chosen. However, as has been argued in the preceding, meaning effect on lexical production is rather limited. The amount of meaning and the type of meaning may be determinant in rather general terms the number of types to be used (it has to be a gross estimation), the basic proportions of content and function words, possibly also a rough range of repetition rates of the word-types to be used. It is to be noted that all those parameters are still subject to sociolinguistic variation. But compared to the choice of particular words, i.e. the actually-produced word-type lists, those variations may be at a different magnitude. They are comparatively small (as shown in Table 1 and Table 2) so that they can be taken as a constant slot for the particular word-type variation. When we can assume that a certain amount of function types are needed for the lexical composition of a stretch of speech and those function type slots can be filled by different function words (slots which must be filled by a particular type are excluded), the fact that those slots were filled with these particular types instead of any other types which are also used in the speech community is a result of sociolinguistic variation. To put things more concretely, words used in speaking are to a large extent like phonological variants in speech in the way that they are functionally equivalent at certain semantic levels. A different phonetic form may not change the referential meaning of a word; likewise, a different lexical type may not change the referential meaning of a discourse.

Since the variation in function types is easier to understand as being insensitive to topic condition, in the following we examine content type variation in the same respect. The shared content types in two speaker-sample matches are listed. In analyzing these lists, we would like to answer the questions: a) Do the content types shared by two speakers result from the fact that they were talking about the same things? b) Do the differences between two social groups in shared content words result from the fact that they had different things to say?

List 1

I. Shared content types by speakers of the same sex

MALE-MALE MATCH	FEMALE-FEMALE MATCH
ban (do)	
	bi (compare)
da (big)	
dao (come to)	dao (come to)
	duo (much)
diao (fall)	
gei (give)	gei (give)
	gongzuo (work)
guo (pass)	
	hao (good)
haoxiang (seem)	
	huiqu (go back)
jia (home)	jia (home)
	jiang (speak)

	jue (feel)
	kan (look)
keyi (may)	keyi (may)
lai (come)	lai (come)
	neng (can)
qu (go)	qu (go)
ren (person)	ren (person)
	renjia (other people)
shang (up)	shang (up)
shihou (time)	shihou (time)
shi (be)	shi (be)
	shi (thing)
shuo (say)	shuo (say)
xialai (downward)	
	xia (down)
	xiao (small)
xiang (think)	xiang (think)
xiang (resemble)	
	yao (want)
you (have)	you (have)
	yuan (far)
	zhao (search)
	zuo (do)
N=20	N=30

II. Shared content types by speakers of same dialect background

BEIJING-BEIJING MATCH	SOUTHERN-SOUTHERN MATCH
	ban (do)
banfa (way)	
	bijiao (compare)
chi (eat)	
da (big)	
dao (come to)	dao (come to)
dixia (ground)	
	dong (understand)
dui (pile)	
	dushu (study)
gan (dare)	
gan (do)	
gei (give)	gei (give)
	gao (do)
	hao (good)
	haoxiang (seem)
jia (home)	jia (home)
	jiang (speak)

jieguo (result)	
	jue (feel)
kan (look)	
keyi (may)	keyi (may)
kou (put upside down)	
lai (come)	lai (come)
li (inside)	
libian (inside)	
	laoshi (teacher)
	nian (year)
	qian (money)
qu (go)	qu (go)
ren (person)	ren (person)
	renjia (other people)
shang (up)	shang (up)
shi (be)	shi (be)
shi (-like)	
shihou (time)	shihou (time)
shuo (say)	shuo (say)
wan (end)	
	xia (down)
	xialai (downward)
xiaoshihou (childhood)	
xiang (resemble)	
xiang (think)	xiang (think)
xiao (small)	
	yao (want)
you (have)	you (have)
	yue (month)
zhan (stand)	
	zuo (do)
(N=29)	(N=27)

Although the discrimination of a particular part of speech for Chinese words taken out of context is difficult and controversial (as has been noted in the footnote 1), judging by their general semantic content, the words on the lists do not appear to be topic-specific. Very few of those content types designate a nominal concept, i.e. things to be talked about; and among the few which do, most of them have generic meanings, such as /ren/ (person), /shihou/ (time). The semantically-identifiable epithets are even fewer and they also have very general meaning, such as /hao/ (good), /xiao/ (small). Most of the types appearing on the lists would probably be used as verbs, judging from their semantic denotation. Interestingly, they are either indicators of some general actions, such as /zuo/ (do), /qu/ (go); or expressing-devices like the English linking-verbs, such as /shi/ (be), /haoxiang/ (seem); or like the assertive verbs in English (e.g. I *think* that...), such as /xiang/ (think), /jue/ (feel); or like the English modals, such as /keyi/ (may), /dei/ (must). In general, the listed shared content types do not have very particular referential meaning. Therefore, it is not apparent that they were used by both speakers out of the necessity of referring to the same specific topic.

Referring to List 1, in contrasting female-shared types with male-shared types, we find that the following were shared by females which were not by males: /bi/ (compare), /gongzuo/ (work), /hao/ (good), /huiqu/ (go back), /jiang/ (speak), /jue/ (feel), /kan/ (look), /neng/ (can), /renjia/ (other

people), /xia/ (down), /xiao/ (small), /yao/ (want), /zhao/ (search), and /zuo/ (do). Does this list indicate that females rather than males tend to "compare", to "look", to "feel", to "go back", to "speak", to "search"; more concerned with "work", and "other people"; more likely to talk about things which are "good" and "small", and associated with "down"? Pending the answer to the above question, here we first assert that it would certainly be far-fetched to conclude that females are thus more concerned with ability (using /neng/ (can) more often), "want" more (sharing /yao/ (want)) and "do" more (sharing /zuo/ (do)). In the preceding it was already made clear that those words all have very general meanings. Therefore, they could have been used to refer to various specific topics. Besides, the categorization of things as being "good" or "bad", "big" or "small" is already a conceptual characterization which is an essential part of the semantic configuration of lexicon, definitely subject to sociolinguistic variation. For /xia/ (down), we have /xialai/ (downward) in the male-shared list, which is not always a synonym of /xia/ (down), but certainly belongs to the same semantic field at the topic level. Therefore, it may be an instance of "alternate ways of saying the same thing". The same can be also said of /jiang/ (speak) (female-shared only) and /shuo/ (say) (both female- and male-shared) alternation, or a /zuo/ (do) and /ban/ (do) alternation. A complete disparity of meaning and form is not advocated here for it is an unrealistic notion. The meaning a speaker intends to express must decide to a certain degree (i.e. in a certain probability) the choice of some words in a discoursal unit. Therefore, it is not necessary to disqualify each of those words as having a close topic-association. However, a closer look on the actual usage of some of the shared words may shed further light on the question. For instance, do our two female speakers "look" more since they both used the word /kan/ (look) which does not appear on the male-shared list? The following are the actual instances of usage of /kan/ in the data by the two female speakers:

a) Southern Female: *Ni dao jie shang ni keyi KAN dianying.*
you come to street up you may look movie
"When you come out, you can watch a movie"

b) Beijing Female: *Ni KAN, hai you jiu na zhong shi*
you look still have just that sort thing
dou neng jizhu.
all can remember
"You see, there are also that kind of things which can be remembered"

c) Beijing Female: *Erqie wo yi KAN jizhe wo dou jue de kepa.*
moreover I one look journalist I feel DE[9] horrible
"Moreover, I feel horrible when I see a journalist"

Not only the word was used referring to different things by different speakers, it was used differently by the same speaker. The same is also true of /zhao/ (search).

d) Southern Female: *Zhi shi yinwei dao xianggang bu dong*
only be because come to H.K. not understand
guangdonghua ZHAO bu dao shiqing.
Cantonese search not come to thing

[9]For explanations of this and other untranslatable terms, see Appendix I.

"Only because (they) came to Hong Kong and did not understand Cantonese thus not being able to find a job"

e) Beijing Female: *Wo jiu bu yuanyi zuo na zhong interview de*
I just not willing do that sort INTERVIEW
shi, ZHAO renjia.
DE thing search other people
"I just don't want to do the kind of thing as contacting people and interviewing them"

f) Beijing Female: *Namo ji shiwan tiao de dongxi*
that way several hundred thousand piece DE thing
wo dao shi chucuo shang nar ZHAO qu ya?
I come to time error up where search to EXCL
"That thing contains several hundred thousands pieces, when it goes wrong, how can I find the error?"

The following content types were shared by the two males but not by the females: /ban/ (do), /da/ (big), /diao/ (fall), /guo/ (pass), /haoxiang/ (seem), /xialai/ (downward), and /xiang/ (resemble). Those are all words of very general and highly abstract meaning except perhaps /diao/ (fall) and /guo/ (pass). One may presume that "something which was falling (or fell down)" could be a specific topic shared by the two speakers, and that "something (someone) passed somewhere" might be an identical part in two similar stories. Looking into the original data, we find that those were not the cases:

g) Beijing Male: *Jiu haoxiang shao le xiang na xianghui*
just seem burn LE insense that insense ash
DIAO xialai le.
fall downward LE
"Just like the fallen insense ash after the insense has burnt off"

h) Southern Male: *Dao disan ge yue, laibuji le: ni fu DIAO*
come to third piece month too late LE you pay fall
le ni hai yi qun jiaxiao.
LE you still one flock family
"When it came to the third month, it worked no longer; if you paid it off, how could you support your family"

i) Beijing Male: *Wo shun zhe nabian GUO.*
I follow ZHE that-side pass
"I passed there"

j) Southern Male: *Wo bu xiaode shang ge libai jiang GUO meiyou?*
I not know up piece week speak pass no
"I wonder if I already talked about it last week"

To summarize, there is no evidence shown that any topics conditioned the dissimilarity in lexical sharedness between the male and female speakers, thus the dissimilarity is best explained by differences in ways of expressing rather than in what is to be expressed.

Referring to List 1, we also find some non-overlapping parts between the two shared-type lists of the dialect-groups. Those non-overlapping parts, which distinguish the dialect-groups, do not appear to be any more topic-dependant than those found with the sex-groups. On the Beijing-Beijing Match list, /chi/ (eat), /dixia/ (ground), /dui/ (pile), /kou/ (put upside down), and /xiaoshihou/ (childhood), judging by the specificity of their semantic contents[10], may possibly be associated with certain topics. But, those types occupy only 17% of the whole list, and 31% of the non-overlapping list (uniquely shared by the group). On the Southern-Southern Match list, /dushu/ (study), /laoshi/ (teacher), and /qian/ (money) may be associated with certain topics. But they take up only 11% of the whole list, and 23% of the non-overlapping list. Therefore, the dialect-groups differ from each other mainly in words not associated with specific topics.

As a summary, both the sex-groups and the dialect-groups differ from each other in the matter of shared content types mainly in words not associated with specific topics. When the large amount of topic-neutral function words are also taken into consideration, one can get the general picture that topic plays an insignificant role in the speakers' overall sharing of word-types.

6. Frequency variation

In the discussion in 5. we treated the types shared by all samples as if they were an invariable part of lexical production. However, a closer look tells us that there is frequency variation among the lexical types, across speakers, and across social groups. Of the words used by all the four speakers, twelve words were used in frequencies of more than 30 times altogether. We divide those twelve words, according to their different frequencies associated with individual speakers, into sexually-biased and dialectally-biased words. The dialectally-biased words are further divided into Beijing-biased words and Southern-biased words, while sexually-biased words are further divided into male-biased words and female-biased words. The percentages of usage frequencies of those biased words are given in Table 6. All the differences in usage frequency between social groups are significant except one. As can be noted, these words cannot be said to be related to any referential meaning. So they are highly indicative of the speakers expressive characteristics. In other words, this means that certain groups of speakers may have certain favourite vocabularies which are excessively employed in their speech.

It is noticeable that there are more dialectally-biased words than sexually-biased words. Besides, it is already noted in 4. that dialect-grouping is more differentiating in sharing ratios than sex-grouping. Moreover, Table 7 shows that, while several differences in frequency distribution between dialect groups are significant, none is significant between sex groups. These facts show that for our data dialect background is a more important factor than sex in conditioning lexical variation. Although it is not my intention to generalize this result to a much wider range, it is conceiveble that dialect (geography) makes a more distinctive division in setting off a locus of intensive linguistic interaction than a cross-dialectal sex division.

[10] Indeed this is a superficial judgement, for instance, one of the speakers used the word /chi/ (eat) in talking about taking medecine, the other used it in the expression /chi-ren/ (man-eating) in a figurative sense.

TABLE 6

1. Beijing-biased word-types

Word-type	Beijing	Southern	Word-token	Difference
wo (I)	69%	31%	71	p<0.05
jiu (just)	71%	29%	68	p<0.05
zai (at)	63%	37%	50	p<0.05
le (LE)	61%	39%	44	p<0.05
Mean	66%	34%	53.5	

2. Southern-biased word-types

Word-type	Southern	Beijing	Word-token	Difference
de (DE)	60%	40%	82	p<0.05
shi (be)	54%	46%	39	N.S.
Mean	57%	43%	60.5	

3. Male-biased word-types

Word-type	Male	Female	Word-token	Difference
ta (she)	79%	21%	43	p<0.05
yi (one)	66%	34%	35	p<0.05
yige (a)	67%	33%	33	p<0.05
Mean	70.7%	29.3%	37	

4. Female-biased word-types

Word-type	Female	Male	Word-token	Difference
bu (not)	29%	29%	41	p<0.05
ni (you)	22%	22%	32	p<0.05
Mean	69.5%	25.5%	36.5	

TABLE 7

Proportion of word-types by frequency

INPUT SAMPLE	NUMBER OF TOKENS PER TYPE										TYPE TOTAL
	1+	2+	3+	4+	5+	6+	7+	8+	9+	10+	
Southern	66%	34%	26%	21%	17%	14%	11%	8%	7%	6%	300
Beijing	44%	29%	22%	14%	11%	8%	7%	7%	6%	4%	321
Difference	*	N.S.	N.S.	*	*	*	N.S.	N.S.	N.S.	N.S.	
Male	47%	31%	24%	17%	13%	11%	8%	6%	5%		311
Female	44%	31%	23%	15%	12%	10%	8%	7%	6%		325
Difference	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.		

* $p < 0.05$

Summary

In a moderate-sized (250 tokens upward) speech corpus, usually many more content types than function types are used. But the large number of content types are used in very low frequencies, while the small number of function types tend to be used in much higher frequencies than the content words. The function-types are used in such high frequencies that function-tokens usually take a larger proportion than content-tokens in the speech corpus. This suggests that the general composition of the lexical-tokens of a speech corpus does not reflect as much what is said in meaning as how it is said (in form). As the most frequently used types tend to be function-types or content-types with very general meaning, they reflect even more ways of speaking than meaning of the speech.

As speech corpus grows large (either by including more speech of a same speaker, or by lumping together speech of different speakers of the same language), the proportion of types used in the corpus keeps growing smaller. This illustrates that either individual vocabularies or the lexicon of a language is rather limited compared to various intentional meanings of real life. In this respect, lexical concepts to discoursal meaning are just like phonemes to phonological words: limited number of items making quite unlimited number of combinations. The differences are that the possible combinations of the former are in a wider scale than the ones of the latter and that the combinations of the former are less conventionalized than the latter thus being more subject to sociolinguistic variation.

Although there is sociolinguistic variation at all levels of lexical variation, the sharing ratios and usage frequencies are qualitatively more apparent than repetition ratios, content and function proportions, type proportion and frequency correlations. The sharing ratios are in nature also indica-

tors of usage frequency, but in a lower scale. The shared types in the samples are not the only ones the speakers could possibly share in their lives, nor can we guarantee that the shared-type list would turn out exactly the same if two different samples of speech were drawn from the recordings by the same speakers. What sharing ratios really tell us is that those speakers all use such types to such a frequency that picking up a sample from the speech of each one, you will probably find those types in all those samples. Therefore, everything boils down to the question of usage frequency. Our findings show that speakers tend to use a specific small group of words more often on all occasions. Those words consist of mostly function words and are semantically neutral to topics. They are used mostly due to force of habit rather than a deliberate choice. So, we can call them a "habitual vocabulary". Interestingly enough, we find that speakers who have similar social attributes also have similar habitual vocabularies.

In a traditional approach of dialectology, words are used as important indicators of dialects. The categorical usage (using or not using) of certain words demarcates dialectal boundaries. In this study, we find that, words are not only categorically, but also quantitatively associated with social groups. In Colby (1966), semantically-categorized words are studied as to their frequencies in folk tales in the purpose of distinguishing cultural patterns. The notion emphasized there, that much can be inferred about the story by the words, is only indirectly touched by a further inference from the content of the story to the interest of the story-teller. In Poplack, Sankoff, & Miller (1987), various borrowed words are found to be associated with different social groups. Since the borrowed words are practically all content words, the association would still be attributed to the interests of the speakers and topics of the speech. However, in this study, we find that many topically-neutral word-types are associated with certain social groups by a high utility frequency. Therefore, the habitual vocabulary, which consists of all such types, is a sociolinguistic indicator.

REFERENCES

COLBY, B. (1966), "Cultural patterns in narrative", dans *Science*, 151, 793-798.

LABOV, W. (1966), *The social stratification of English in New York*, Washington, D.C., Center for Applied Linguistics.

LABOV, W. (1972), *Sociolinguistic patterns*, Philadelphia, Unversity of Pennsylvania Press.

LABOV, W. (1982), "Building on empirical foundation", dans Lehmann, W. et Y. Malkiel (eds.), *Perspectives on historical linguistics*, Amsterdam, Benjamins, pp. 17-92.

LAVANDERA, B. (1978), "Where does the sociolinguistic variable stop?", dans *Language in society*, 7, pp. 171-182.

POPLACK, S., SANKOFF, D. et MILLER, C. (1987), *The social correlates and linguistic consequences of lexical borrowing and assimilation*, University of Ottawa.

SANKOFF, D. (1986), "Sociolinguistic and syntactic variation", dans F. Newmeyer (ed.), *Linguistics: the Cambridge survey*, New York, Cambridge University Press.

APPENDIX I

A Glossary

BA: preposition, causative.

DAO: verbal complement, indicating that there is a result caused by the act referred by the verb.

DE: modifier marker; nominalisation marker; when used after a predicate verb and before an NP, emphasizing the agent of the verb, or the time or location; when used at the end of a statement, intensifying the affirmative mood; after a verb, expressing capability or possibility; after a verb or an adjective, introducing a resultative or degree complement.

EXCL: exclamatory markers, expressing a variety of moods, or expressing no meaning but being used only as rhythmic fillers.

LE: aspectual marker, expressing the completion of an action or change of state; when used after a verb, expressing affirmative, pressing, or dissuasive moods, or indicating the appearance of a new condition.

ONOM: onomatopoeia.

ZHE: aspectual marker, progressive.

APPENDIX II

Type-proportion and frequency distribution correlation in two samples of an English speaker's speech[11]

INPUT SAMPLE	NUMBER OF TOKENS PER TYPE										TYPE TOTAL
Topic samples	1+	2+	3+	4+	5+	6+	7+	8+	9+	10+	
Claire's Ego	31%	13%	9%	6%	2%		1%			1%	140
Car Accident	35%	17%	13%	10%	8%	6%	5%	2%		1%	119

[11]The two speech samples were supplied by Prof. Shana Poplack.

MARQUEURS DE DISCOURS ET FONCTION DISCURSIVE DES NARRATIONS[1]

Charleen Rains
Université de Montréal

1. Introduction

Les marqueurs de discours sont des mots ou expressions qui jouent un rôle structural ou interactif au niveau discursif dans la langue orale. En position initiale d'énoncé ces mots marquent le lien, continuité ou rupture, entre l'énoncé à suivre et le discours en cours, ou entre le locuteur et l'interlocuteur (Vicher et Sankoff, 1986; Schiffrin, 1987, par exemple). Les marqueurs ne sont qu'une des ressources employées afin de souligner la cohérence dans le discours. Jefferson (1978) présente plusieurs moyens par lesquels les participants à une conversation peuvent intégrer leurs contributions. Dans la recherche présentée ici, j'étudie la présence de marqueurs en début de narration et entre les diverses phrases à l'intérieur de la narration. L'étude porte sur la distribution de ces marqueurs par rapport à la structure narrative et à la fonction discursive de la narration, dans le but de comprendre le rôle de ces marqueurs dans l'organisation du discours. Je présenterai d'abord les données, puis la méthodologie et finalement, les résultats.

2. Les données

Les narrations étudiées ici sont tirées des entrevues enregistrées du corpus de Montréal 1984 (Thibault, Vincent, Sankoff et Kemp). Les locuteurs interviewés sont tous des Montréalais francophones. L'analyse porte sur 46 narrations produites par 13 locuteurs. Ces narrations sont, soit sollicitées par l'intervieweur, soit spontanées, c'est-à-dire présentées comme arguments ou exemples intégrés dans le discours du locuteur, ou comme anecdotes enchaînées dans la suite (ou glissement) des idées de la conversation.

Plus particulièrement, j'ai examiné les mots et expressions extra-syntaxiques en position initiale d'énoncé.

(1) "FAIT QUE LA les enfants ont commencé à dire....."
(2) "PUIS elle avait elle avait fait du brocoli."
(3) "BIEN Beaconsfield a appelé à la SPCA."

3. La méthodologie

3.1. Les parties d'une narration

Inspirée des critères de Labov (1972), j'ai dégagé la structure de chaque narration. Un locuteur peut commencer sa narration avec un résumé (ou abstract), par exemple, "Je me souviens quand j'ai mis des jeans pour la première fois." Il peut la terminer avec une coda, qui ramène la conversation au présent, "Puis ça c'est un exemple que je donne comme ça là tu sais." Toutefois le résumé et la

[1]Cette recherche a été effectuée grâce à une bourse du Conseil de recherche en sciences humaines du Canada.

coda ne sont pas toujours présents dans les narrations. L'histoire comme telle se trouve dans les énoncés narratifs qui racontent les événements en ordre chronologique. Ces énoncés sont, en général, accompagnés par d'autres énoncés dont la fonction est d'évaluer l'histoire (énoncé d'évaluation) ou de l'orienter quant aux personnages, au temps, au lieu (énoncé d'orientation).

3.2. Les segments narratifs et non narratifs

Dans cette étude, les énoncés d'évaluation et d'orientation sont regroupés sous l'étiquette "non narratif". L'un et l'autre peuvent être insérés dans la narration des événements et présenter une rupture dans la séquence narrative. Ces ruptures, en conjonction avec les reprises de la séquence narrative, séparent une narration en segments narratifs, où l'on trouve le récit des événements et, en segments non narratifs, où l'on trouve les évaluations personnelles du locuteur ainsi que des informations supplémentaires. Un segment peut contenir un ou plusieurs énoncés.

3.3. L'énoncé

Les énoncés sont définis selon des critères syntaxiques et prosodiques. Dans l'exemple (4), l'intonation finale suivie d'une pause, ainsi que le ton plus élevé du début de l'exemple (5), fragmentent en deux ce qui aurait pu être un seul énoncé.

(4) "L'an passé je suis allé au Centre Epic là.
(5) Faire de des exercices physiques."

Dans l'exemple (6) le locuteur se reprend, soit pour se corriger, soit pour préciser le pronom "il" et ce, sans pause, en conservant la même intonation. Il s'agit d'un seul énoncé.

(6) "FAIT QUE il s'est adressé la fille s'est adressé à moi."

Cependant, dans les exemples (7) à (9), le locuteur interrompt son énoncé afin d'ajouter une information. Ensuite, il reprend le fil de son histoire.

(7) "Il a même refait du
(8) Mon père est un gars qui avait une grosse volon une grande volonté.
(9) Il est même allé faire du ski par après."

Dans la perspective de cette recherche, qui est d'analyser la présence des marqueurs en position initiale, cet exemple représente trois énoncés dont un, le (7), est incomplet. Voir aussi les exemples (10) à (15).

(10) "Tout à coup on m'env
(11) LA je travaille plus.
(12) On m'env
(13) J'ai aucune j'avais pas de pension."

(14) "DONC LA on s'est prom
(15) TU SAIS on a fait quand même un asses bon but."

Toutes les narrations sont ainsi analysées en segments narratifs et non narratifs dans lesquels on retrouve des énoncés complets et des énoncés incomplets dont la fonction (narrative ou non) est évidente. Les énoncés incomplets et non identifiables sont nécessairement exclus de ces segments.

3.4 La fonction discursive des narrations

Raconter une expérience personnelle est un fait commun dans une conversation. Dans une entrevue, la narration n'est pas toujours sollicitée par l'intervieweur. Elle peut être produite spontanément par l'interviewé à l'intérieur d'une argumentation ou d'une explication pour valider ou justifier un propos. De plus, par une association d'idées ou par glissement de sujet, l'interviewé peut être porté à raconter une anecdote qui lui vient à l'esprit, et ce, sans qu'il n'y ait d'autre fonction discursive que la communication amicale d'une histoire intéressante. Dans mes données, j'ai distingué ces trois types de narrations:

1) **celles qui servent d'exemple ou d'argument;**
2) **celles qui sont sollicitées par l'intervieweur et servent de réponse;**
3) **celles qui sont racontées sans autre but évident que leur intérêt propre comme anecdote.**

3.5 L'analyse

J'ai d'abord étudié le début des narrations, c'est-à-dire les mots et expressions reliant ces narrations à la conversation. Ensuite, j'ai examiné la distribution des marqueurs selon la fonction des énoncés (narrative ou non) dans la structure narrative pour chaque type de narration. Enfin j'ai quantifié les marqueurs en position de transition (au début d'un segment) et de séquence (à l'intérieur d'un segment), et cela toujours selon la fonction de l'énoncé et le type de narration.

4. Les résultats

4.1 Les débuts d'une narration

Le tableau 1 présente les mots et les expressions qui ouvrent les narrations.

TABLEAU 1

Les débuts d'une narration

MARQUEURS TEMPORELS	MARQUEURS PRÉ-PHRASTIQUES	CONNECTEURS	AVERTISSEMENTS	AUCUN
je me souviens(2)	ah	mais (3)	ça c'est comme	
l'autre fois	bien (5)	parce que (2)		
l'an passé (2)	oui	puis (2)	la pire qui	
une journée		et	m'est arrivé	
la dernière fois		et puis (2)	m'as te dire	
un moment donné		alors	une affaire	
après ça		comme (6)		
il y a une fois				
Total: 10	7	17	3	9

(Les chiffres entre parenthèses indiquent le nombre d'occurrences pour les marqueurs qui apparaissent plus d'une fois.)

On voit que 10 des 46 narrations débutent par un marqueur temporel, une variation du cliché "il était une fois". Neuf de ces marqueurs commencent une narration pendant un tour de parole de l'interviewé. Le dixième enchaîne une narration immédiatement après une question de l'intervieweur.

Sept des narrations commencent avec ce que Vicher et Sankoff (1986) appellent les "pre-sentential markers" (traduit ici par "marqueurs pré-phrastiques"). Tous ces marqueurs ouvrent une narration qui suit directement une question de l'intervieweur, sauf une exception où la narration succède à une réponse du locuteur quelque peu contradictoire.

(16) "Je suis abonné à Dimanche Matin mais je l'aime pas Dimanche Matin.
Non, je l'aime pas Dimanche Matin.
BIEN c'est une c'est une femme qui m'a téléphoné ici. PUIS..."

Les connecteurs enchaînent 17 narrations dans le discours du locuteur déjà en cours; aucun de ces connecteurs ne commence à la fois un tour de parole et une narration.

Trois narrations commencent par des expressions qui annoncent en quelque sorte la nature du discours à suivre. "Ça c'est comme..." et "Le pire qui m'est arrivé..." font un lien explicite avec le discours précédent. "M'as te dire une affaire" met l'interlocuteur en garde d'être bien attentif à ce qui suit. Dans ces trois cas, l'expression sert à la fois de disjonction dans le discours en cours du locuteur, et de liaison entre la narration à suivre (une narration-exemple) et le discours précédent.

Enfin, 9 narrations commencent sans mot ou expression marquant explicitement la transition vers un discours narratif. Il est intéressant de noter qu'aucune de ces narrations n'est une narration-anecdote; ces dernières sont toutes introduites par un marqueur qui souligne la transition.

En fait 35 narrations, soit 76%, sont enchaînées dans la conversation avec un mot ou une expression. Parmi ces 35 marqueurs, 24 (ou 69%) sont de simples connecteurs ou des marqueurs pré-phrastiques employés dans l'enchaînement des phrases à l'intérieur des narrations. Ces mots servent à marquer des liens à plus d'un niveau du discours.

4.2 Les marqueurs dans la structure narrative

Le tableau 2 présente le nombre et la fréquence des énoncés narratifs et non narratifs qui commencent par un marqueur. Ces marqueurs peuvent être soit des connecteurs, soit des "marqueurs pré-phrastiques", soit des explétifs (MON DIEU, OUFF, AYE, etc.). Il y a aussi quelques exemples de marqueurs interactifs (TU SAIS, OK?).

TABLEAU 2

Les énoncés marqués selon leur fonction et le type de narration

Type de narration	Fonction de l'énoncé narrative	non narrative	total
exemple	40,7% 33/81	49,6% 62/125	46,1% 95/206
réponse	59,1% 39/66	50,0% 58/118	53,3% 97/182
anecdote	48,3% 28/58	48,3% 34/78	45,6% 62/136
total	48,8% 100/205	48,3% 154/319	48,5% 254/524

On constate qu'en général, il y a autant d'énoncés narratifs que non narratifs qui sont "marqués" (48%). Toutefois, dans la catégorie "exemple", les énoncés narratifs sont moins souvent marqués que les énoncés non narratifs (40% contre 49%), ce qui contraste avec la catégorie "réponse" (59% contre 50%).

On voit également une corrélation entre le type d'énoncé plus souvent marqué et la fonction discursive de la narration. La catégorie "réponse" regroupe les narrations qui suivent une variante de la question "Qu'est-ce qui est arrivé?". Ce qui est arrivé (les événements) se trouve dans les énoncés narratifs: ils sont plus souvent marqués. Dans les narrations-exemples, le poids de l'exemple comme argument à l'intérieur d'un discours se trouve dans les énoncés d'orientation et d'évaluation - ici ce sont eux qui sont le plus souvent marqués. Le peu de différence de fréquence de marqueurs pour les deux catégories d'énoncés dans les narrations-anecdotes serait donc dû au fait qu'ici, ni les événements, ni les évaluations ne dominent.

On remarque enfin que la proportion totale d'énoncés marqués est plus importante pour la catégorie "réponse" (53% contre 45-46%). On se souvient que les narrations de cette catégorie sont sollicitées par l'intervieweur, tandis que les narrations des deux autres catégories sont produites spontanément par l'interviewé. L'analyse de la variance[2], un facteur à la fois, démontre que la différence de fréquence entre les narrations-réponses (sollicitées) et les narrations-exemples/anecdotes (spontanées) est significative.

4.3. Les marqueurs intersegmentaux et les marqueurs intrasegmentaux

Comme on l'a vu plus haut, les narrations sont constituées de segments narratifs et non narratifs. Les marqueurs INTERsegmentaux sont les marqueurs qui commencent un segment; ils marquent une transition, une rupture ou une reprise de la séquence narrative. Les marqueurs INTRAsegmentaux se trouvent à l'intérieur d'un segment; ils marquent le lien entre les énoncés du même type (narratif ou non). Dans le Tableau 3, on voit la distribution des marqueurs dans ces deux positions. Le total est moindre que dans le Tableau 2 car certains énoncés, ceux qui suivent immédiatement un énoncé incomplet et non identifiable, sont exclus.

TABLEAU 3

Énoncés marqués en position inter- et intrasegmentale par type de narration

Énoncé	Exemple	Réponse	Anecdote
intersegment			
narratif (reprise)	44,7% 17/38	64,7% 22/34	44,4% 12/27
non narr. (rupture)	46,3% 19/41	60,5% 23/38	46,7% 14/30
total	45,6% 36/79	62,5% 45/72	45,6% 26/57
intrasegment			
narratif	37,2% 16/43	51,6% 16/31	45,2% 14/31
non narratif	51,2% 43/84	46,0% 35/76	42,2% 19/45
total	46,5% 59/127	47,7% 51/107	43,4% 33/76

[2]Je tiens à remercier Pierrette Thibault pour l'analyse de la variance et pour ses commentaires sur le texte. Je remercie également Diane Vincent pour son aide et encouragement ainsi que mes collègues Michelle Daveluy et Carole Laurin pour leurs conseils.

Ce tableau nous démontre que les marqueurs intersegmentaux sont plus fréquents dans la catégorie "réponse" que dans les deux autres catégories (62% contre 45%). Ceci signifie que les ruptures, comme les reprises de la séquence narrative, sont plus fréquemment accompagnées de marqueurs dans une narration qui sert de réponse à une question. On ne remarque pas une telle différence dans la distribution des marqueurs intrasegmentaux (48% contre 46 et 43%). Les catégories narrative et non narrative ne démontrent pas de différence significative dans la distribution des marqueurs.

5. Conclusion

Une analyse plus sophistiquée ferait sans doute ressortir le poids des individus et de leurs stratégies dans les calculs généraux. Néanmoins, il semble que la présence des marqueurs soit liée à la fonction de la narration dans le discours. On les trouve plus souvent dans les narrations sollicitées que dans les narrations spontanées, produites dans le but d'apporter un exemple comme argument ou tout simplement de raconter une histoire intéressante. De plus, dans les narrations sollicitées, on les trouve plus fréquemment entre les énoncés de fonctions différentes (narratives ou non), qu'entre les énoncés ayant la même fonction. On ne remarque pas une telle différence dans les narrations spontanées.

Bien sûr, il faudra élargir ce corpus afin de confirmer les hypothèses présentées ici. Il faudra également regarder de plus près les mots choisis pour ce rôle d'enchaînement afin de voir si leurs caractéristiques sémantiques usuelles influencent le type d'enchaînement qu'ils marquent.

BIBLIOGRAPHIE

JEFFERSON, G. (1978), "Sequential aspects of storytelling in conversation", dans Jim Schenkein (dir.), *Studies in the organization of conversational interaction*, New York, Academic Press.

LABOV, W. (1972), *Language in the inner city*, Philadelphia, University of Pennsylvania Press.

SCHIFFRIN, D. (1987), *Discourse markers*, New York, Cambridge University Press.

VICHER, A. et SANKOFF, D. (1986), "The emergent syntax of pre-sentential turn openings", Communication présentée à NWAV-XV, Stanford University.

VARIATION PHONÉTIQUE DANS L'EMPLOI DES PRONOMS DE TROISIÈME PERSONNE EN FRANÇAIS QUÉBÉCOIS

Marie-Josée Bourget
Université Laval

1. Introduction

Pour notre communication, nous présenterons l'élision des pronoms de troisième personne comme un cas de variation phonétique. Le corpus que nous utilisons est celui décrit dans *Le français parlé dans la ville de Québec: une étude sociolinguistique*. Il comporte des enregistrements d'adolescents de Sainte-Foy et de Saint-Sauveur qui ont été interviewés seuls ou en groupe ainsi que des interviews de certains de leurs parents. De cet échantillon comportant 54 jeunes et 42 de leurs parents, un sous-échantillon a été constitué pour l'étude des pronoms personnels.

En prenant connaissance du tableau 1, on peut constater que ce corpus comprend 53 locuteurs: 10 adolescents de 13 à 16 ans à Sainte-Foy et le même nombre à Saint-Sauveur dont 5 garçons et 5 filles dans chacun des groupes. Un ou deux de leurs parents ont été interviewés pour un nombre total de 16 à Sainte-Foy et 17 à Saint-Sauveur.

Nous avons choisi ce sous-échantillon pour pouvoir vérifier les différences et/ou similitudes d'utilisation des pronoms selon l'âge, le sexe, le quartier de résidence, la profession ainsi que le degré de scolarité.

Les interviews à l'intérieur de notre sous-échantillon ont toutes été transcrites et c'est à partir de ces transcriptions et des bandes sonores que notre étude s'est effectuée. Toutes les variantes de certains pronoms de 3ème personne, soit *il*, *ils*, *elle*, *elles*, *le*, *la*, *les*, *en*, *leur*, *lui* et *eux*, ont été relevées. Ceci implique tout autant les variantes standard que les variantes non standard.

2. Présentation des variables

Un premier relevé de tous les pronoms personnels avait déjà été effectué sur feuille de disposition de données sur cartes indiquant les variables sociales des locuteurs interviewés, soit le numéro de code, l'âge, le sexe, le quartier de résidence, le temps de résidence dans le quartier, la mobilité géographique, le nombre d'enfants dans la famille, le rang de l'informateur, le type d'habitation et le type d'entrevue. Également, la profession et le nombre d'années de scolarité du père, du mari ou de l'informateur, dépendamment de la personne interrogée, ont été indiqués ainsi que la profession et le nombre d'années de scolarité de la mère, de l'épouse ou de l'informatrice.

En plus des données sociales de l'informateur, les informations suivantes ont été indiquées: le thème de discussion en cause, le procédé d'élicitation, c'est-à-dire les éléments utilisés par l'interviewer pour susciter une réponse de l'interviewé (ex.: question posée avec *tu*, *vous*, etc..), le fait qu'il s'agisse d'un discours du locuteur ou d'un discours rapporté, le pronom employé, son référent, sa fonction syntaxique, sa forme phonétique, sa place par rapport au verbe, le verbe lui-même, sa forme verbale, la personne du verbe, le type de phrase et, enfin, la nature et la fonction du syntagme dans lequel le pronom a été employé. Toutes ces données ont été saisies sur ordinateur et constituent un premier fichier de données SAS (STATISTICAL ANALYSIS SYSTEM) à partir duquel nous n'avons

retenu que les pronoms de troisième personne. Nous avons réécouté toutes les bandes sonores et vérifié tous les pronoms à analyser. Nous avons apporté les corrections nécessaires au fichier brut à l'aide du système MUSIC.

TABLEAU 1:

Répartition des 53 locuteurs en termes de sexe d'âge et quartier de résidence

	SAINTE-FOY	SAINT-SAUVEUR
ADOLESCENTES	001	028
	002	041
	010	045
	036	052
	048	055
ADOLESCENTS	005	004
	016	014
	021	026
	032	046
	044	056
PARENTS	[1]101-201	104-204
	102-202	114-214
	105-205	126-226
	110	128
	116-216	141
	121	145-245
	132-232	146-246
	136	152
	144	155
	148-248	156
		[2]356-456

Cependant, comme cet ensemble de données ne suffisait pas à l'élaboration de notre analyse, nous avons ajouté les informations suivantes: le contexte phonétique précédant le pronom, c'est-à-dire s'il est précédé d'une consonne, d'une voyelle, d'une pause et même, du [z] ou du [t] de liaison. Ces mêmes critères à l'exception du [z] ou du [t] de liaison ont servi à codifier le contexte suivant le pronom. Nous avons également indiqué si le pronom était à l'initiale absolue de l'énoncé, à l'intérieur ou en finale absolue. Ces trois variables ont été codifiées mais n'ont malheureusement pas pu faire partie de notre analyse puisque leur saisie n'est pas encore terminée.

[1]Les chiffres de la première centaine indiquent qu'il s'agit des mères alors que les pères se sont vu assigner un chiffre de la deuxième centaine.

[2]Ces numéros correspondent aux grands-parents du locuteur 056.

3. Variables sociales et linguistiques

Ainsi, nous prendrons en considération deux catégories de variables: sociologiques et linguistiques.

3.1 *Variables sociales*: Les variables sociales qui seront prises en considération pour cette analyse seront l'âge, le sexe, le quartier de résidence, la profession et le niveau de scolarité.

a) *variable âge*: La variable âge ne comporte que deux catégories: adolescents (13 à 17 ans) et parents (35 à 65 ans).

b) *variable quartier de résidence*: Les quartiers de résidence sont Saint-Sauveur, quartier de classe ouvrière et Sainte-Foy, quartier de classe moyenne.

c) *variable profession*: La variable profession a été divisée en 10 catégories.

1. OUVRIER NON SPÉCIALISÉ
(débardeur, chauffeur, messager, livreur, femme de ménage, repasseuse, aide boulangère)

2. OUVRIER SPÉCIALISÉ
(mécanicien, infirmière auxiliaire, etc...)

3. CONTREMAÎTRE
(contremaître, petit commerçant)

4. SECRÉTAIRE
(secrétaire, agent de bureau, commis, agent d'immeuble, donne des cours de yoga, etc...)

5. VENDEUR
(vendeur, serveur, caissier, coiffeur, etc...)

6. GÉRANT
(gérant, directeur des ventes, etc...)

7. PROFESSEUR
(professeur, infirmière, spécialiste en informatique, etc...)

8. PROFESSIONNELS
(professionnel, directeur général, etc...)

9. MÉNAGÈRE,RETRAITÉ
(ménagère, retraité, etc..)

0. CHÔMEUR
(chômeur, prisonnier, donnée manquante)

d) *variable scolarité*: La variable scolarité a été divisée comme suit:

1. 1 à 6 années d'études inclusivement correspondent au niveau primaire terminé ou non terminé.

2. 7 à 12 années de scolarité inclusivement correspondent au niveau secondaire terminé ou non terminé.

3. 13 et 14 années de scolarité correspondent au rang collégial terminé ou non terminé.

4. 15 années de scolarité et plus correspondent au niveau universitaire terminé ou non terminé.

3.2 *Variables linguistiques*: Nos variables linguistiques sont d'ordre syntaxique et lexical.

a) *variables syntaxiques*:
1. La fonction syntaxique du pronom.
2. La place du pronom dans une séquence.
3. Le verbe avec lequel il est en rapport.

b) *variable lexicale*: 1. Le référent du pronom.

4. Élision des pronoms de troisième personne

À l'intérieur de notre analyse, nous avons trouvé que seulement trois pronoms pouvaient être complètement élidés: *il*, *ils* et *elles*. Nous présenterons pour chacun les contextes linguistiques et sociaux dans lesquels on peut les retrouver, c'est-à-dire que nous vérifierons si l'emploi des variantes non standard (ici l'élision du pronom) est plus élevé à Saint-Sauveur qu'à Sainte-Foy, chez les hommes que chez les femmes, chez les adolescents que chez les parents, chez les non-professionnels que chez les professionnels et chez ceux qui ont moins de scolarité versus ceux qui en ont plus et si cet emploi des variantes est également lié à des contraintes d'ordre linguistique. Ceci représente en fait nos hypothèses de base.

4.1 *Élision du pronom IL*

Nous avons tout d'abord vérifié la distribution de fréquence de l'élision et de la non-élision du pronom *il*. Il y a, en tout, 11 844 pronoms *il* à l'intérieur de notre corpus. 7 401 (62,5%) n'ont pas été élidés et 4 442 (37,5%) ont été élidés. Nous avons également constaté que cette élision se faisait généralement avec les verbes *sembler* et *falloir*: "il me semble" est souvent dit "me semble" et "il faut que", "faut que".

TABLEAU 2:

Élision du pronom *il* en fonction de son référent

RÉFÉRENT	NON ÉLISION	ÉLISION
SANS RÉFÉRENT	20,5% (1 136)	79,5% (4 417)
DÉL.[3](E) DÉT.	99,6% (5 130)	0,4% (23)
DÉL.(E) IND.	99,7% (1 138)	0,3% (3)

En considérant le tableau 2, on peut constater que le pronom *il* est essentiellement élidé lorsqu'il n'a pas de référent et à un taux assez élevé (79,5%). Lorsque le référent représente un(e) délocuté(e) déterminé(e) ou non, les pourcentages d'élision sont trop minimes pour en tenir compte (0,3 et 0,4%) et sont probablement explicables en termes d'erreur.

TABLEAU 3:

Élision du pronom *il* en fonction de l'âge des informateurs

ÂGE	NON ÉLISION	ÉLISION
ADOLESCENTS	64,5% (2 835)	35,5% (1 560)
PARENTS	61,2% (4 544)	38,8% (2 880)

Le tableau 3 démontre que les parents élident un peu plus le pronom *il* que les adolescents: 38,8% contre 35,5%. L'indice de probabilité p pour cette relation est inférieur à 0,05, ce qui signifie que la relation est significative.

Ce résultat va à l'encontre de nos hypothèses de base. Nous croyions au départ que les adolescents élideraient davantage que leurs parents. Une différence importante est pourtant à noter. Le nombre absolu de pronoms utilisés par les adolescents est de 1 560 contre 2 880, presque la moitié moins. Peut-être les parents portent-ils moins attention à leur discours étant donné qu'ils parlent plus? Nous l'ignorons mais nous vérifierons pour d'autres variations phonétiques si les résultats confirment ou infirment ces dires.

[3]DÉL: délocuté; DÉT: déterminé; IND:indéterminé.

TABLEAU 4:

Élision du pronom *il* en fonction du sexe des locuteurs

SEXE	NON ÉLISION	ÉLISION
FÉMININ	61,7% (4 613)	38,3% (2 867)
MASCULIN	63,8% (2 766)	36,2% (1 573)

À partir du tableau 4, on constate que les femmes élident davantage que les hommes: 38,3% contre 36,2%. On s'attendait à ce que ce soit l'inverse étant donné que plusieurs études sociolinguistiques arrivent à la conclusion que les femmes utilisent davantage les formes prestigieuses que les hommes.

Il faut toutefois noter que les femmes utilisent 7 480 pronoms *il* tandis que les hommes en utilisent 3 339. La raison de cette différence de comportement est peut-être explicable par rapport à cette situation.

TABLEAU 5:

Élision du pronom *il* en fonction du quartier de résidence des locuteurs

QUARTIER DE RÉSIDENCE	NON ÉLISION	ÉLISION
SAINT-SAUVEUR	61,3% (3 209)	38,7% (2 025)
SAINTE-FOY	63,3% (4 170)	36,7% (2 415)

En prenant connaissance du tableau 6, nous constatons que l'élision du pronom *il* est moins fréquente à Sainte-Foy (36,7%) qu'à Saint-Sauveur (38,7%). Étant donné que l'indice de probabilité p est inférieur à 0,05, nous pouvons parler de relation significative et devons conclure que la relation quartier de résidence versus élision du *il* ne peut être due au hasard.

4.1.1 Élision du pronom *il* pour les femmes

TABLEAU 6:

Élision du pronom *il* en fonction de la scolarité des femmes

SCOLARITÉ	NON ÉLISION	ÉLISION
PRIMAIRE	65,8% (1 094)	34,2% (568)
SECONDAIRE	56,7% (1 019)	43,3% (779)
COLLÉGIAL	64,6% (328)	35,4% (180)

Nous voyons d'après le tableau 7 que la scolarité des femmes ne comporte que 3 catégories c'est-à-dire qu'il n'y a pas de femmes qui sont allées à l'université dans notre échantillon. Les femmes qui élident le plus sont celles du niveau primaire (65,8%) mais celles du niveau collégial (64,6%) ont presque le même taux. Le nombre d'occurences de pronom *il* pour les femmes du niveau de scolarité collégial est de 508 pronoms contre 1 662 pour les femmes du niveau primaire. Peut-être ces dernières portent-elles moins attention à leur discours?

TABLEAU 7:

Élision du pronom *il* en fonction de la profession des femmes

PROFESSION	NON ÉLISION	ÉLISION
OUVR. NON SPÉC.	58,6% (163)	41,4% (115)
SECRÉTAIRE	39,3% (126)	60,7% (195)
VENDEUSE	0% (0)	100% (4)
PROFESSEURE	63,8% (367)	36,2% (208)
MÉNAGÈRE	64,9% (2 070)	35,1% (1 118)

Le tableau 8 est très difficile à commenter puisque les femmes sont en grande majorité des ménagères et le nombre d'occurences des pronoms pour cette catégorie est vraiment très élevé comparativement aux autres catégories: 3 188 pour les ménagères, 4 pour les vendeuses, 278 pour les

ouvrières non spécialisées, 321 pour la catégorie secrétaire, et 575 pour les professeures. Peut-être faudrait-il vérifier la scolarité de ces femmes ménagères pour réussir à tirer une conclusion de ces faits.

4.1.2 Élision du pronom *il* pour les hommes

TABLEAU 8:

Élision du pronom *il* en fonction de la scolarité des hommes

SCOLARITÉ	NON ÉLISION	ÉLISION
PRIMAIRE	65,5% (403)	4,5% (212)
SECONDAIRE	56,8% (708)	43,2% (539)
COLLÉGIAL	64,2% (556)	35,8% (310)
UNIVERSITÉ	45,8% (151)	54,2% (179)

L'élision du pronom *il* pour les hommes de niveau universitaire est moins élevée que pour les hommes des autres niveaux comme nous nous y attendions. Toutefois, on ne peut pas dire que l'élision du pronom *il* est inversement proportionnelle au niveau de scolarité des hommes puisque le comportement des hommes des niveaux 3 et 1 est pratiquement du même ordre: 64,2% contre 65,5%.

TABLEAU 9:

Élision du pronom *il* en fonction de la profession des hommes

PROFESSION	NON ÉLISION	ÉLISION
OUVR. NON SPÉC.	57,2% (305)	42,8% (228)
OUVR. SPÉC.	59,3% (256)	40,7% (176)
CONTREMAÎTRE	65,0% (327)	35,0% (176)
GÉRANT	53,4% (287)	46,6% (250)
PROFESSEUR	53,0% (300)	47,0% (266)
RETRAITÉ	70,4% (343)	29,6% (144)

L'ordre d'élision du pronom *il* selon la profession des hommes est la suivante: catégories professeur (47%), gérant (46,6%), ouvrier non spécialisé (42,8%), ouvrier spécialisé (40,7%), contremaître (35%) et finalement, retraité (29,6%). Nous partions de l'hypothèse que les retraités élideraient le plus et la catégorie professeur le moins et nous arrivons tout à fait au contraire. Nous ne savons comment expliquer cette situation sinon qu'il aurait peut-être mieux valu classer les retraités selon le type le type d'emploi qu'ils occupaient avant de prendre leur retraite.

4.2 *Élision des pronoms ils et elles*

Comme l'emploi de ces deux pronoms est relativement du même ordre, nous avons pensé les commenter ensemble. Toutefois, le nombre d'occurences de chacun est différent: le pronom *ils* est utilisé 5 407 fois et il est élidé dans 6,6% des cas alors que le pronom *elles* est utilisé 65 fois et est élidé dans 29,2% des cas. Pour les deux pronoms, la relation référent versus élision du pronom est non pertinente ($p>0,05$). Le pronom *ils* occupe la position sujet seulement alors que le pronom *elles* occupe les fonctions sujet et prépositionnelle et n'est jamais élidé pour cette deuxième fonction. Ces deux pronoms sont élidés lorsqu'ils sont employés avec le verbe ou l'auxiliaire *être*.

TABLEAU 10:

Élision des pronoms *ils* et *elles* en fonction de l'âge des locuteurs

ÂGE		NON ÉLISION	ÉLISION
ADOLESCENTS	ILS	91,5% (1 628)	8,5% (152)
	ELLES	15,4% (4)	84,6% (22)
PARENTS	ILS	94,3% (3 414)	5,7% (206)
	ELLES	38,5% (15)	61,5% (24)

L'examen du tableau 10 nous permet de constater que l'élision des pronoms *ils* et *elles* est plus importante chez les adolescents que chez leurs parents: 8,5% contre 5,7% pour le pronom *ils* et 84,6% contre 61,5% pour le pronom *elles*. Par contre, on peut remarquer que le pronom *ils* est moins souvent élidé que le pronom *elles* autant pour les jeunes que pour les adultes: 8,5% d'élision de *ils* chez les adolescents, 5,7% chez les parents contre 84,6% d'élision de *elles* pour la première catégorie et 61,5% pour la seconde. Il est difficile d'expliquer ces différences mais le nombre d'occurences pour chacun de ces pronoms peut y être pour quelque chose. Pour le pronom *elles* surtout, comme 25% des cases comportent moins de 5 occurences, il se peut que le test du chi carré ne soit pas valide.

Les relations de ces pronoms versus le sexe des locuteurs n'ont pas été retenues car l'indice de probabilité était supérieur à 5%.

TABLEAU 11:

Élision du pronom *ils* en fonction du quartier de résidence des informateurs

QUARTIER DE RÉSIDENCE		NON ÉLISION	ÉLISION
SAINT-SAUVEUR	ILS	92,5% (2 505)	7,5% (204)
SAINTE-FOY	ILS	94,3% (2 537)	5,7% (154)

Nous n'avons pu retenir les relations incluant le pronom *elles* puisque l'indice de probabilité était égal à 27,5%. D'après ce tableau, l'élision du *ils* est plus fréquente à Saint-Sauveur (7,5%) qu'à Sainte-Foy (5,7%) comme nous nous y attendions. Le taux d'élision n'est pas très élevé mais n'est pas à notre avis négligeable.

4.2.1 Élision des pronoms *ils* et *elles* pour les femmes

L'élision des pronoms *ils* et *elles* pour les femmes ne peut être mise en relation avec leur scolarité puisque l'indice de probabilité est supérieur à 5%.

En ce qui a trait aux professions des femmes, nous pouvons retenir seulement que la catégorie ménagère comporte les locutrices qui élident le plus mais également qui utilisent ces deux pronoms plus que les autres. Nous ne voyons pas l'utilité de présenter les résultats puisque l'indice de probabilité est trop élevé pour cette relation.

4.2.2 Élision des pronoms *ils* et *elles* pour les hommes

Nous ne traiterons pas de l'élision du pronom *elles* pour les hommes puisque l'indice de probabilité de ce pronom en fonction de leur scolarité est de 12,2% et en fonction de leur profession 27,0%.

TABLEAU 12:

Élision du pronom *ils* en fonction de la scolarité des hommes

SCOLARITÉ	NON ÉLISION	ÉLISION
PRIMAIRE	278 (90,9%)	28 (9,1%)
SECONDAIRE	711 (95,9%)	30 (4,1%)
COLLÉGIAL	189 (93,6%)	13 (6,4%)
UNIVERSITÉ	165 (98,8%)	2 (1,2%)

Le tableau 12 démontre que l'élision du pronom *ils* est inversement proportionnelle à la scolarité des hommes c'est-à-dire que plus les hommes sont scolarisés, moins ils élident le pronom en question. Ceci confirme vraiment ce à quoi nous nous attendions.

TABLEAU 13:

Élision du pronom *ils* en fonction de la profession des hommes

PROFESSION	NON ÉLISION	ÉLISION
OUVR. NON SPÉC.	361 (91,6%)	33 (8,4%)
OUVR. SPÉC.	178 (93,7%)	12 (6,3%)
CONTREMAÎTRE	92 (93,9%)	6 (6,1%)
GÉRANT	205 (96,2%)	8 (3,8%)
PROFESSEUR	303 (97,4%)	8 (2,6%)
RETRAITÉ	204 (97,1%)	6 (2,9%)

La relation élision du pronom *ils* en fonction de la profession des hommes représente une relation très forte puisque l'indice de probabilité est de 0,6%. Ces deux variables sont inversement proportionnelles comme nous l'avions prévu au départ. Toutefois les catégories professeur ainsi que retraité sont du même ordre: non élision de 97,2% ± 0,2% et élision de 2,7% ± 0,2%. Le taux d'élision de ce pronom passe de 8,4% à 2,9% en partant de la catégorie ouvrier non spécialisé jusqu'à celle de retraité.

5. Conclusion

En conclusion, nous pouvons affirmer que l'emploi des variantes standard et non standard, c'est-à-dire l'emploi du pronom sous quelque forme que ce soit versus son non-emploi, n'est pas du même ordre à Saint-Sauveur et à Sainte-Foy, chez les hommes et chez les femmes, chez les adolescents et chez les parents, chez les non-professionnels et chez les professionnels, chez ceux qui ont moins de scolarité versus ceux qui en ont plus. L'élision du pronom *il* versus celle du pronom *ils* est presqu'entièrement contradictoire, sociologiquement parlant.

Notre corpus comporte également des contraintes d'ordre linguistique pour l'emploi de la variante non standard analysée: l'élision du pronom. Pour le pronom *il*, on sait qu'il est élidé seulement s'il est sans référent et employé avec les verbes *falloir* et *sembler* surtout[4]. Pour les pronoms *ils* et *elles*, leur élision se produit en fonction sujet et avec le verbe ou auxiliaire *être*.

Le comportement de l'élision du pronom *il* infirme nos hypothèses de base tandis que celui des pronoms *ils* et *elles*, lorsque les relations sont pertinentes, les confirme.

[4]Il est également élidé avec le verbe avoir dans la séquence "il y a" mais nous n'avons pu traiter de ce phénomène lors de notre communication.

BIBLIOGRAPHIE

DESHAIES, D. (1981), *Le français parlé dans la ville de Québec: une étude sociolinguistique*, Publié par le Centre International de Recherche sur le Bilinguisme, Publication G-1, Québec, 86 pages.

LABOV, William (1978), *Le parler ordinaire*, Les Éditions de minuit, Paris.

L'ALTERNANCE ENTRE LES DÉTERMINANTS DÉMONSTRATIFS [Sə], [Sɛt] ET [STə] EN FRANÇAIS PARLÉ À MONTRÉAL[1]

Michelle Daveluy
Université de Montréal

Dans le cadre d'une recherche de maîtrise, j'ai analysé la variation dans l'usage des déterminants démonstratifs en français parlé à Montréal et ce, en temps réel.

C'est l'analyse des facteurs linguistiques qui influencent la prononciation [stə] qui est décrite ici. Pour bien situer cette recherche, une brève présentation des différentes formes des déterminants démonstratifs en français parlé et des catégories d'analyse utilisées est nécessaire. Par la suite, on verra comment les variantes sont regroupées pour respecter leurs contraintes d'occurrence. Enfin, on indiquera que les explications fournies à ce jour pour rendre compte de l'emploi de la forme [stə] ne s'excluent pas les unes des autres. Ainsi, et l'affaiblissement de la valeur déictique des déterminants démonstratifs et l'assimilation articulatoire ont un effet sur les variations considérées bien qu'une hypothèse de neutralisation morphologique de l'opposition de genre soit au coeur de cette recherche.

A. Les déterminants démonstratifs du singulier en français

En ancien français, il existe deux séries de démonstratifs: ceux de la forme CIST qui correspondent à la notion sémantique de proximité et ceux de la forme CIL qui se rapportent à l'éloignement.

Au cours du XIV^{e} siècle, une distinction fonctionnelle apparaît entre ces deux séries de démonstratifs de sorte qu'il faut discriminer les adjectifs démonstratifs des pronoms démonstratifs. Les premiers sont issus de la série en CIST tandis que les seconds proviennent de la série en CIL.

À peu près à la même époque, l'opposition sémantique entre les deux séries de démonstratifs s'affaiblit. En effet, entre le XIII^{e} et le XV^{e} siècles, apparaissent les particules CI et LA, qui sont adjointes aux démonstratifs, et ce sont elles qui marquent maintenant l'opposition entre ce qui est proche et ce qui est éloigné. On aboutit au modèle qui suit.

1. Cette maison ci/Cette maison là

En français moderne, donc, les déterminants démonstratifs du singulier sont CE, CET pour le masculin et CETTE pour le féminin.

---

[1]Je tiens à remercier ma directrice de maîtrise, Madame Pierrette Thibault, qui m'a guidée tout au long de cette recherche ainsi que Monsieur David Sankoff qui m'a beaucoup aidée pour l'analyse statistique des données. Merci aussi à Carole Laurin pour son support lors de la préparation de cette communication et à Charleen Rains qui a lu ce texte.

B. Les déterminants démonstratifs du singulier en français parlé

Déjà on peut noter que les déterminants démonstratifs CET et CETTE ont une prononciation identique. En français parlé, plusieurs auteurs mentionnent aussi la prononciation [st ə] dans les cas où ces déterminants sont attendus.

Ainsi, dès le XVII[e] siècle, cette prononciation est relevée par Thomas Corneille, poète et écrivain français, qui souligne qu'"il eut été pédant de prononcer autrement" (Brunot et Bruneau, 1969: 144). Dans Brunot et Bruneau (1969: 201), aux remarques de Corneille s'ajoutent celles de Vaugelas.

> "'Dans le discours familier, on prononce ST' homme, STE femme, dit Thomas Corneille, et ce serait une affectation vicieuse de dire CET homme, CETTE femme.' Même 'd'excellents Prédicateurs', en chaire, prononçaient: 'ST' action, ST' habitude' (Vaugelas, Remarques, Chassang, II,164)."

Pourtant, selon Brunot et Bruneau (1969: 201), "cette manière de parler est aujourd'hui populaire et dialectale."

D'autre part, Henri Bauche (1946: 89) note que la forme [st ə] est aussi utilisée dans des contextes où le déterminant démonstratif CE est attendu. L'exemple qu'il donne est celui de CTE CHEVAL.

L'extension de la prononciation [st ə] au masculin devant consonne fait disparaître l'opposition de genre à l'oral puisqu'il n'y a aucune distinction entre [ST ə] FILLE et [ST ə] GARS.

En français québécois, Léard (1978) remarque des tendances analogues à celles décrites pour le français de France. Il établit un parallèle entre la prononciation [st ə], devant voyelle, et l'emploi de l'article élidé *l'* devant une voyelle. Cette analogie fait cependant abstraction de l'utilisation de la forme [st ə] devant une consonne tant au féminin qu'au masculin bien que cet auteur considère que l'explication des formes du féminin éclaircira le cas de la variante [st ə] au masculin. Cette variante [st ə] au masculin devant consonne est d'ailleurs attestée par Seutin dans ses travaux sur le français parlé à l'île aux Coudres (1975: 116).

> 2.1 ...vu qu'on avait entendu C'TE moteur là.
>
> 2.2 On mettrait un drap blanc sur C'TE tréteau là.
>
> 2.3 Pis, la fille, alle était pris avec C'TE grand énergumène là.

Enfin, une autre source de neutralisation de l'opposition de genre en français québécois est mentionnée par Fournier (1981: 49). Il s'agit de la prononciation [s][2] au féminin, comme dans cet exemple qui a été relevé dans mes données.

---

[2]Dans cette analyse, la chute du [ə] n'a pas été considérée comme un critère déterminant une variante. Ainsi, [s] et [st] ne sont pas discriminés de [s] et [st]. C'est la chute ou l'insertion d'un [t] qui a été l'objet de l'analyse.

3. Je le manquerai pas CETTE [s] fois là. (1.84.3250)[3]

À la lumière des différents compte-rendus sur les prononciations des déterminants démonstratifs du singulier, il apparaît donc que la variation est possible entre [sə], [sɛt] et [stə], dépendamment du contexte d'occurrence.[4]

C. Les catégories d'analyse

Conformément à l'approche de la sociolinguistique, la distribution des variantes de chaque variable[5] est analysée selon les facteurs linguistiques et sociaux qui peuvent l'influencer.[6]

Trois catégories de facteurs d'ordre linguistique ont été considérées. Il s'agit, en premier lieu, d'éléments du contexte d'occurrence qui ont trait à la valeur déictique associée aux déterminants démonstratifs. Dans une deuxième catégorie, on retrouve des groupes de facteurs qui se rapportent aux caractéristiques sémantiques des noms que les démonstratifs déterminent ainsi qu'aux propriétés discursives du contexte. Enfin, les particularités de l'environnement phonique des occurrences sont regroupées dans une troisième catégorie.

Chacune de ces catégories ou dimensions de l'analyse est constituée de différents facteurs regroupés en sous-catégories. Ainsi, la valeur déictique associée aux déterminants démonstratifs est analysée en considérant l'effet de l'adjonction d'une particule au déterminant démonstratif et l'influence du rôle thématique du groupe démonstratif. Les caractéristiques sémantiques et discursives des énoncés analysés contiennent trois groupes de facteurs: la référence sémantique du mot déterminé par le démonstratif, la thématique du discours et l'effet discursif associé à l'énoncé dans lequel le démonstratif est produit. Enfin, l'environnement phonique des occurrences comporte deux groupes de facteurs qui permettent de vérifier l'influence de l'assimilation articulatoire sur la prononciation [stə]. Il s'agit du contexte phonétique de droite et de la présence d'un *t* sonore dans le contexte d'occurrence, comme dans:

4. Bien tu as pas vu ça ce que j'ai FAIT CETTE nuit.
[fɛtstə] (121.3070)

---

[3]Le premier chiffre correspond au numéro du locuteur; le deuxième, à l'année de l'entrevue (pour les locuteurs interviewés à deux reprises); enfin, le troisième chiffre indique la ligne dans les transcriptions de l'entrevue (pour 1984, ces numéros de ligne sont temporaires puisque l'aménagement du corpus n'était pas complété lors de cette recherche).

[4]La prononciation [sta] pour le féminin, notée par Fournier (1981:49) et attestée par Lafolette (1969:54), n'a pas été relevée dans les corpus de données utilisés pour la recherche rapportée ici.

[5]Pour une définition des notions de variables et variantes en sociolinguistique, voir Labov, 1976: 129-130.

[6]Les catégories d'ordre social qui ont été utilisées sont au nombre de trois: le sexe des locuteurs, le groupe d'âge auquel ils appartiennent ainsi que leur position sociale. Sans présenter en détail l'analyse sociale, notons simplement que ces trois catégories influencent la variation entre les formes [sə], [sɛt] et [stə]. Dans l'emploi de la variante [stə] il n'y a pas évidence d'un changement en cours. Il s'agit plutôt d'un phénomène relié à la gradation d'âge. En ce qui concerne la variante [sə] au féminin, on observe les caractéristiques d'un changement en cours bien que le peu de données ne permette pas d'analyse statistique significative.

Puisque la valeur déictique associée aux déterminants démonstratifs constitue la particularité de l'analyse, il faut expliquer en détail en quoi elle consiste.

D. La valeur déictique associée aux déterminants démonstratifs

En soi, la valeur déictique caractérise différents éléments du discours qui se définissent seulement par rapport à l'instance du discours où ils sont produits (Benveniste,1966: 262). Par exemple, un énoncé dans lequel CE LIVRE est désigné ne prend son sens qu'en fonction du locuteur qui émet l'énoncé et ce, au moment et dans la situation où il le fait.

Cependant, avec l'apparition des particules CI et LA entre le XIII[e] et le XV[e] siècles, il est généralement considéré que cette valeur déictique associée aux déterminants démonstratifs a diminué. En effet, la notion sémantique de proximité et d'éloignement est maintenant désignée par ces particules adjointes aux démonstratifs.

En français québécois du XX[e] siècle, Fournier (1981: 46) considère que le système des démonstratifs est caractérisé par une phase de synthétisation des formes qui a pour effet de créer une forme figée pour les adjectifs démonstratifs. Il postule que la particule CI est remplacée par LA et que le renforcement doit se faire à l'aide des adverbes ICI et LA afin de conserver l'opposition entre ce qui est proche et ce qui est éloigné. D'après lui, il est donc impossible de trouver ICI comme premier élément adjoint au démonstratif et l'absence de la particule LA rend la séquence agrammaticale. Pourtant, ces deux cas sont relevés dans mes données, bien que leur nombre soit restreint.

5.1 Puis là bien...de CE [sə] temps ICI ça a été la soirée du hockey. (7.71.521)

5.2 C'est pas la question que c'est plus CE [stə] quartier ICI. C'est juste parce que je travaille ici. (66.84.4520)

5.3 Ils ont eu une bourse pour faire CE [sə] travail? (75.71.633)

5.4 ...quand je voyais CETTE [st] émission je fermais la télévision. (88.84.13150)

En ce qui concerne l'adjonction de la particule ICI comme premier élément, j'aurais tendance à nuancer les propos de Fournier. Pour le français de France, Frei mentionne que cette particule tend à disparaître tandis que Bauche (1946: 90) signale que CI se change souvent en ICI. Il serait probablement plus juste de parler d'utilisation peu fréquente plutôt que d'une impossibilité pure et simple.

D'autre part, l'agrammaticalité des phrases provoquée par l'absence de la particule LA me semble discutable. L'exemple apporté par Fournier est le suivant:

6. * Prends cette assiette [prɑ̃stasyɛt]

Cet auteur remarque qu'il n'y a que deux cas, en français québécois populaire, où un démonstratif peut être utilisé sans particule. Il s'agit des expressions temporelles marquant la continuité avec le moment présent ou encore d'interjections et de tours figés. Voici les exemples avec lesquels il illustre ces cas.

7.1 **cet été, ce printemps, cet automne, cet hiver, cette année,...**
[stete, sprɛ̃tɑ̃, stotɔn, stiv ɾ, stane]

7.2 **bien voyons donc, cette affaire! [stafaɛ:r]**

cette histoire [stistwaɛ:r] de toujours arriver à une heure pareille!

En français parlé à Montréal, comme on l'a vu en 5.3 et 5.4, il existe pourtant des cas autres que ceux-ci où un démonstratif est utilisé sans la particule LA ou encore avec CI.

8. **Surtout dans ' Pointe il y en avait deux (théâtres). Un sur CE [s] bord CI, un l'autre bord. (51.84.1020)**

De toute façon, il est à noter qu'avec ou sans particule, la prononciation des déterminants démonstratifs varie:

9.1 **Bien je vas sûrement jouer au tennis aussi CET [s t] été dans le fond là. Mais je me propose de jouer au golf CET [st] été. (13.84.1520)**

9.2 **Là c'est tranquille depuis CE [sə] temps là. (15.84.2670)**
...puis j'ai pas travaillé CE [stə] temps là. (15.84.4610)

D'ailleurs, certains auteurs accordent une valeur sociale aux différentes particules. Pour Bruneau et Brunot (1969: 206), ICI serait associé au langage populaire tandis que, selon Frei (1929: 86), l'opposition entre (I)CI et LA ne correspond plus à la valeur sémantique de proximité ou d'éloignement, mais plutôt à une opposition stylistique entre le relevé et le populaire.

D'autre part, toujours dans l'hypothèse de l'affaiblissement de la valeur déictique des déterminants démonstratifs, il est possible que le rôle thématique du groupe démonstratif ait aussi une influence sur la distribution des différentes variantes d'une variable.

Le rôle thématique consiste en ce dont il est question dans un énoncé, le sujet du discours, en opposition à ce qui est dit sur ce sujet (commentaire, rhème, propos, prédicat). La topicalisation est d'ailleurs un procédé linguistique qui consiste à faire d'un constituant de la phrase, le topique, c'est-à-dire le thème, dont le reste de la phrase est le commentaire.

Après une première analyse, il est apparu que certains groupes de facteurs reliés à la structure de l'énoncé ont un effet d'interaction avec la fonction du groupe démonstratif (sujet et complément prépositionnel ou non). Par exemple, l'apposition ou l'extraposition dans un énoncé favorisent la topicalisation. C'est dans ce sens que Brunot (1953: 281) parle de la mise en lumière des sujets ou des idées dans des tours comme: CETTE FEMME, ELLE...

Donc, pour vérifier l'effet du rôle thématique du groupe démonstratif, sept facteurs ont été retenus. Pour les sujets, trois cas sont possibles dépendamment du fait qu'ils soient apposés ou non.

10.1 **apposé et non extraposé**
J'ai décidé d'aller à St-Henri aussi parce que CETTE [st] école là elle finit en secondaire trois. (131.2760)

10.2 apposé et extraposé vers la droite
Il a été battu CE [stə] petit gars là. (2.71.540)

10.3 non apposé et non extraposé
CE [s] capital là reviendra peut-être plus ici. (76.71.271)

Pour les compléments, il y a quatre possibilités basées sur la présence d'une préposition et sur l'extraposition.

11.1 prépositionnel extraposé vers la gauche
Depuis CE [sə] temps là, Lorraine je l'ai jamais revue. (7.71.492)

11.2 prépositionnel non extraposé
Tu mets de CE [st] savon là, tu le mets dans la chaudière...(90.84.420)

11.3 sans préposition extraposé vers la gauche
Puis CET [sɛt] hiver j'ai travaillé un peu. (125.2570)

11.4 sans préposition non extraposé
Fait-qu'on aime mieux CE [sə] journal là que les autres. (25.84.13910)

E. Les corpus de données

Cette étude sur les déterminants démonstratifs s'inscrit dans le cadre d'un vaste projet de recherche sur le français parlé à Montréal. Les donnnées analysées proviennent de deux corpus: le corpus Sankoff-Cedergren et le corpus Montréal 1984.

Le premier a été recueilli en 1971 par une équipe dirigée par Gillian Sankoff, Henrietta Cedergren et David Sankoff. Ce corpus est composé de 120 entrevues réalisées auprès de francophones d'origine montréalaise qui sont âgés d'au moins 15 ans. Cet échantillon respecte la diversité aux niveaux géographique et économique et en ce qui concerne l'âge et le sexe des locuteurs.

En 1984, c'est sous la direction principale de Pierrette Thibault et Diane Vincent qu'un nouveau corpus de 72 entrevues a été constitué. Soixante de ces entrevues consistent en une deuxième rencontre avec des personnes qui avaient été interviewées en 1971. Les 12 autres forment un corpus de jeunes locuteurs, âgés entre 15 et 25 ans, qui a été ajouté afin d'assurer la comparabilité des données. Les critères utilisés à l'origine pour le corpus Sankoff-Cedergren ont été respectés au moment du recueil de ce corpus de jeunes locuteurs.

L'objectif premier visé lors de l'élaboration de ces deux corpus est le même: recueillir des conversations le plus près possible du style informel malgré les contraintes inhérentes à la situation d'entrevue. La durée moyenne des entrevues de 1971 est d'une heure environ tandis qu'en 1984 elle se situe aux alentours d'une heure trente.

Pour la recherche sur les déterminants démonstratifs, 132 entrevues ont été dépouillées, soit 60 en 1971 et 72 en 1984. Au total, 2152 occurrences de déterminants démonstratifs du singulier ont été relevées.

F. L'analyse statistique

L'analyse statistique des données a été effectuée au moyen de VARBRUL, conçu par David Sankoff. Plus précisément, c'est la version 2.3, adaptée par Suzan Pintzuk pour les micro-ordinateurs IBM, qui a été utilisée.

Ce programme d'analyse produit un modèle probabiliste. Il établit la vraisemblance maximale qu'un corpus de données ait pu être engendré par le modèle selon des valeurs accordées aux facteurs contextuels relevés.

Pour établir la configuration des groupes de facteurs qui est la plus appropriée pour un corpus de données, il faut tenir compte de la signification statistique de l'effet individuel des facteurs dans ces groupes. Le programme VARBRUL arrive à cette fin par une analyse de régression multiple dans laquelle les effets individuels et combinés des groupes de facteurs sont calculés.

L'analyse de régression multiple permet de hiérarchiser les facteurs non déterministes, c'est-à-dire, ceux qui n'influencent pas l'emploi d'une variante ou d'une autre de façon catégorique.

Concrètement, une mesure statistique qui reflète l'influence sur la variante sélectionnée en tant que valeur d'application est déterminée pour chacun des facteurs des groupes retenus. Lorsque ces poids relatifs se situent aux extrémités de l'échelle, soit 0.00 et 1.00, l'effet du facteur est déterminant. Il est donc retiré des étapes ultérieures de l'analyse statistique.

Lors de la dernière itération du programme, les valeurs qui ont un poids relatif se rapprochant de 1.00 ou de 0.00 témoignent d'un effet marqué du facteur. Autour de 0.50, le facteur a peu d'effet.

Dans mes analyses, c'est la variante [st] qui représente la valeur d'application, sauf dans le cas de la variation entre [sɛt], [stə] et [sə] au féminin devant consonne où la variante [sə] joue ce rôle. Les poids relatifs près de 1.00 favorisent donc ces variantes.

G. Le regroupement des variantes

On a vu (en B) que les contextes d'occurrence déterminent, au départ, la forme attendue du déterminant démonstratif selon le genre du mot déterminé et la présence d'une voyelle ou d'une consonne au début de ce mot. Par le fait même, les variantes possibles dans un contexte donné sont restreintes.

Les exemples qui suivent représentent les cas types d'occurrence des prononciations relevées dans les corpus.

12.1 Masculin

Devant consonne

1. **Il était bon CE [sə] gâteau là. (30.71.291)**
2. **On s'est arrangé pour avoir un party CE [stə] soir là. (126.4460)**

Devant voyelle

1. **Mais lui il a gardé CE?ɛ[s t] esprit là. (56.71.394)**
2. **Fait-que tout CET [st] argent là ça a été sur ses garçons. (72.84.1580)**

12.2 **Féminin**

Devant consonne

1. **CETTE [sɛ t] fonction là on en a pas besoin. (56.84.18620)**
2. **Fait-que j'avais acheté CETTE [stə] revue là. (54.84.9570)**
3. **Eux autres ils mangeaient sur CETTE [sə] table là les grands. (2.71.300)**

Devant voyelle

1. **Au début j'étais vraiment réceptive à CETTE [sɛ t] idée là. (7.84.8340)**
2. **Moi j'ai CETTE [st] impression là. (7.71.888)**

Devant voyelle, seulement les déterminants démonstratifs CET et CETTE peuvent être utilisés. La prononciation de ces formes varie entre [sɛt] et [stə].

Devant consonne, CE ou CETTE sont possibles, mais pour un genre spécifique à chacun. Au masculin, CE peut être prononcé soit [sə], soit [stə] tandis que, pour le féminin, CETTE peut varier entre [sɛt], [stə] et [sə].

Ces conditions d'occurrence restreignant les choix possibles dans les déterminants démonstratifs, trois types de variation doivent être considérés. Il s'agit de la variation entre les formes [sɛt] et [stə], entre [sə] et [stə] et, enfin, entre [sɛt], [stə] et [sə].

Pour la variation entre [sɛt] et [stə], deux regroupements de variantes sont étudiés: les occurrences devant voyelle, d'une part, et les cas au féminin devant consonne, d'autre part. Les deux autres variations ne comportent qu'un seul regroupement de variantes chacune (les occurrences du masculin devant consonne et les occurrences du féminin devant consonne respectivement).

H. L'analyse des facteurs linguistiques

C'est en comparant l'effet des facteurs dans les différents regroupements de variantes (quatre au total) qu'on peut analyser les trois variations possibles dans l'emploi des déterminants démonstratifs du singulier.

Ainsi, les tableaux 1, 2 et 3 démontrent des résultats différents dans les variations entre: [sɛt] et [stə] (devant voyelle, tableau 1, et au féminin devant consonne, tableau 2) et [sə] et [stə] (au masculin devant consonne, tableau 3).

Le seul groupe de facteurs statistiquement significatif qui se retrouve dans les trois analyses est celui qui a trait à la référence sémantique du mot déterminé par le démonstratif. Dans chaque cas, c'est la référence concrète (à une personne ou à un objet) qui favorise la variante [stə].

TABLEAU 1

Facteurs linguistiques qui influencent la variation entre [sɛ t] et [stə] devant voyelle

Données de 1971 et de 1984

Groupes et facteurs		Poids relatifs en faveur de [stə]	Occurrences des déterminants démonstratifs		
			[stə]	total	% de [stə]
Particule					
- là		.65	231	330	70
- ci		.39	2	5	40
- ø		.31	118	330	36
Rôle thématique					
- sujet apposé	à droite	.74	40	47	85
- compl. + prép.	non extraposé	.57	94	137	69
- compl. - prép.	non extraposé	.52	114	222	51
- compl. + prép.	à gauche	.41	22	41	54
- compl. - prép.	à gauche	.40	56	166	34
- sujet non apposé	non extraposé	.27	4	11	36
- sujet apposé	non extraposé	.23	5	11	45
Référence sémantique					
- concrète		.62	108	143	76
- locative		.49	220	465	47
- abstraite		.28	23	57	40
Thématique du discours					
- ø		.69	7	9	78
- scolarité		.62	69	118	58
- goûts et activités		.52	162	312	52
- occupation		.42	62	128	48
- résidence		.39	24	49	49
- langue		.36	27	49	55
Total		input .55	351	665	53

TABLEAU 2

Facteurs linguistiques qui influencent la variation entre [sɛt] et [stə] au féminin devant consonne

Données de 1971 et de 1984

Groupes et facteurs	Poids relatifs en faveur de [stə]	Occurrence des déterminants démonstratifs		
		[stə]	total	% de [stə]
Particule				
- là	.67	97	233	42
- ci	.59	3	11	27
- ∅	.10	3	81	4
Effet discursif				
- insistance en conclusion	.93	18	21	86
- description d'un fait	.68	10	21	48
- évaluation	.46	42	150	28
- ∅	.43	26	104	25
- discours rapporté	.36	7	29	24
Référence sémantique				
- concrète	.60	60	141	43
- locative	.46	31	115	27
- abstraite	.36	12	69	17
Total	input.25	103	325	32

TABLEAU 3

Facteurs linguistiques qui influencent la variation entre [s ə] et [stə] au masculin devant consonne. Données de 1971 et de 1984

Groupes et facteurs		Poids relatifs en faveur de [stə]	Occurrences des déterminants démonstratifs [stə]	total	% de [stə]
<u>Rôle thématique du groupe démonstratif</u>					
- sujet apposé	non extraposé	.90	20	36	56
- compl. - prép.	à gauche	.87	13	37	35
- compl. - prép.	non extraposé	.77	61	259	24
- sujet apposé	à droite	.74	16	44	36
- sujet non apposé	non extraposé	.74	5	25	20
- compl. + prép.	non extraposé	.50	38	813	5
- compl. + prép.	à gauche	.30	11	649	2
<u>Référence sémantique</u>					
- ∅		.72	16	94	17
- concrète		.72	100	300	33
- abstraite		.47	15	236	6
- locative		.44	61	1465	4
<u>Effet discursif</u>					
- insistance en conclusion		.95	30	43	70
- description d'un fait		.65	19	115	17
- discours rapporté		.62	22	76	29
- évaluation		.53	85	1038	8
- ∅		.39	36	823	4
<u>Thématique du discours</u>					
- occupation		.64	48	387	12
- goûts et activités		.54	88	834	11
- langue		.44	26	256	10
- résidence		.40	18	354	5
- ∅		.38	2	24	8
- scolarité		.34	10	240	4
<u>Contexte phonétique de droite</u>					
- consonne occlusive		.57	125	1224	10
- consonne fricative		.41	67	867	8
<u>Total</u>		input .05	192	2095	9

Par contre, le groupe de facteurs sur la particule adjointe au déterminant démonstratif est retenu uniquement pour la variation entre [sɛt]et [stə](tableaux 1 et 2). C'est, dans les deux regroupements de variantes utilisés, la particule LA qui favorise la variante [stə]. Ce fait démontre que ce type de variation n'est pas uniquement influencé par la réduction vocalique comme Léard le propose dans son parallèle entre l'article élidé *l'* et cette variante [stə].

D'autre part, la tendance à croire que la production de la forme [stə] au masculin devant consonne (tableau 3) nécessite l'adjonction de la particule LA est infirmée par l'analyse statistique puisque l'influence de ce groupe de facteurs n'est pas significative.

D'ailleurs, il est intéressant de noter que cette variante est aussi influencée de façon significative par des éléments qui ont trait à l'assimilation articulatoire, la présence d'une consonne occlusive au début du mot déterminé par le démonstratif jouant en sa faveur.

C'est donc dire que la variante [stə] dans les différents contextes ne peut être expliquée en fonction d'une seule hypothèse, soit l'affaiblissement de la valeur déictique ou l'assimilation articulatoire, mais plutôt par des phénomènes relevant de ces deux possibilités.

TABLEAU 4

Effet de la particule *là*
dans la variation entre [sɛt], [stə] et [sə]
au féminin devant consonne

Données de 1971 et de 1984

Groupes et facteurs	Occurrences des déterminants démonstratifs		
	[sə]	total	% de [sə]
Particule			
- là	62	295	21
- ∅	1	82	1
- ci	0	11	0

En ce qui concerne la variation au féminin devant consonne ([sɛt], [stə] et [sə], tableau IV), le peu d'occurrences relevées restreint l'analyse statistique bien qu'une tendance très claire ressorte en ce qui a trait à l'adjonction de la particule LA. En effet, 62 occurrences sur un total de 63 sont utilisées avec cette particule. Dans ce contexte, il semble donc y avoir co-occurrence entre la variante [sə] et la particule LA.

I. Conclusion

La présentation des résultats obtenus dans ma recherche a permis de démontrer, premièrement, que la variante [st], autant au masculin devant consonne que dans les autres contextes d'occurrence, est linguistiquement intégrée dans la communauté francophone de Montréal.

D'ailleurs, en ce qui concerne les occurrences au masculin devant consonne, j'ai noté, sans toutefois creuser la question, que le nombre d'items lexicaux auxquels s'applique cette variante a nettement augmenté. En 1971, les 72 occurrences de [stə] au masculin devant consonne totalisent 28 items lexicaux. En 1984, les 120 occurrences dans ce contexte touchent 48 lexèmes, dont 35 qui n'étaient pas utilisés avec [stə] en 1971.

De plus, l'analyse a clairement établi que les hypothèses avancées jusqu'à maintenant pour expliquer la forme [stə] dans les déterminants démonstratifs ne sont pas suffisantes. L'affaiblissement de la valeur déictique aussi bien que l'assimilation articulatoire influencent cette forme, mais il semble qu'une neutralisation morphologique de l'opposition de genre constitue l'explication la plus adéquate. C'est ce qu'indique l'emploi des trois formes de déterminants démonstratifs ([sə], [sɛt] et [stə]) au masculin ET au féminin. La variation dans la prononciation des déterminants démonstratifs en français parlé à Montréal s'inscrit donc dans une tendance générale, en français québécois, à la neutralisation morphologique du genre et du nombre (voir Sankoff,G. et Cedergren, H.,1976; Lemieux et al,1985; Ouellet, 1986). Un apport majeur de cette recherche aura été de démontrer que cette tendance s'applique aussi à la catégorie grammaticale des déterminants.

BIBLIOGRAPHIE

BAUCHE, H. (1946), *Le langage populaire. Grammaire, syntaxe et dictionnaire du français tel qu'on le parle dans le peuple avec tous les termes d'argot usuel*, Paris, Payot.

BENVENISTE, E. (1966), *Problèmes de linguistique générale I*, Paris, Gallimard, Chap. V.

BRUNOT, F. (1926,1953), *La pensée et la langue. Méthode, principes et plan d'une théorie nouvelle du langage appliquée au français*, Paris, Masson et Cie.

BRUNOT F. et BRUNEAU, C. (1969), *Précis de grammaire historique de la langue française*, Paris, Masson et Cie.

DEES, A. (1971), *Étude sur l'évolution des démonstratifs en ancien et en moyen français*, Groningen, Solters-Noordhoff Publishing.

FOURNIER, R. (1981), "Les démonstratifs ... et ça continue" dans *Recherches linguistiques à Montréal*, Vol. 17, pp. 43-56.

FREI, H. (1929,1971), *La grammaire des fautes. Introduction à la linguistique fonctionnelle. Assimilation et différenciation. Brièveté et invariabilité. Expressivité*, Genève, Slatkine Reprints.

LABOV, W. (1976), *Sociolinguistique*, Paris, Editions de Minuit.

LAFOLETTE, J.E. (1969), *Étude linguistique de quatre contes folkloriques du Canada français. Morphologie et syntaxe*, Québec, Presses de l'Université Laval, Les archives de folklore 9.

LEARD, J.M. (1978), "Essai d'interprétation de quelques faits de morphologie du québécois", dans Boisvert et al., *Travaux de linguistique québécoise 2*, Québec, Presses de l'Université Laval, pp. 121-142.

LEMIEUX, M. SAINT-AMOUR, M. et SANKOFF, D. (1985), "TUT en français de Montréal: un cas de neutralisation morphologique", dans Lemieux et Cedergren (éds.), *Les tendances dynamiques du français parlé à Montréal*, Tome 2, Québec, Gouvernement du Québec, pp. 7-90.

OUELLET, M. (1986), *Des participes passés en français parlé à Montréal*, Mémoire de maîtrise non publié, Département d'anthropologie, Université de Montréal.

SANKOFF, D. (sous presse), "Variable Rules", dans Ulrich Ammon, Norbert Dittmar, Klauss J. Mattheier (éds.) *Sociolinguistics. An international handbook of the science of language and society.*

SANKOFF, D. et SANKOFF, G. (1973), "Sample survey methods and computer assisted analysis in the study of grammatical variation", dans R. Darnell (dir.), *Canadian languages in their social context*, Edmonton, Linguistic Research Inc., pp. 7-61.

SANKOFF, G. et Cedergren, H. (1976), "Les contraintes linguistiques et sociales de l'élision du /l/ chez les Montréalais", dans M. Boudreault et F. W. Moehoem (dirs), *Actes du XIIIe congrès international de linguistique et de philologie romanes*, Québec, Presses de l'Université Laval, pp. 1101-1116.

SEUTIN, E. (1975), *Description grammaticale du parler de l'Ile-aux-Coudres*, Montréal, Presses de l'Université de Montréal.

THIBAULT, P. VINCENT, D. et RAINS, C. (1986), "Le corpus Sankoff-Cedergren 13 ans plus tard", Communication présentée au congrès de l'association Canadienne-Française pour l'Avancement des Sciences, ACFAS, Montréal.

À LA RECHERCHE DES DIPHTONGUES NASALES DU FRANÇAIS PARLÉ À QUÉBEC.

Sylvie Dubois
Université Laval

1. Introduction

Lorsqu'on tente d'imiter le parler populaire de Québec, le recours aux diphtongues de certaines voyelles brèves, orales ou nasales est sans doute le phénomène le plus employé et le plus efficace.

ex: L'été, on mange du pain, le matin.

La présente étude porte sur la diphtongaison de trois voyelles nasales en français parlé à Québec: [ɑ̃], [ɛ̃], [ɔ̃]. Pour chacune de ces trois variables, on relève trois variantes: une variante standard, une autre possédant un allongement de la nasalité et une troisième variante impliquant un changement assez important de la qualité du timbre jusqu'à pouvoir lui attribuer le nom de diphtongue[1].

Plusieurs questions se posent au sujet de cette diversité linguistique. La recense-t-on pour toutes les voyelles nasales ou seulement pour certaines d'entre elles? Quels sont les facteurs linguistiques qui influencent ou non le développement, le maintien ou la disparition des variantes? Existe-t-il des différences dans la production des voyelles nasales entre les groupes sociaux et selon une gradation d'âge? En d'autres mots, quel rôle joue le contexte social dans la production des variantes dans le français parlé de Québec[2]?

1.1 *La problèmatique*

Malgré la saillance de la variante diphtonguée à l'extérieur et dans la communauté linguistique de Québec, ce phénomène n'a suscité que peu de travaux systématiques. La littérature est peu abondante sur le sujet et il n'existe pas encore d'étude sur le phénomène des diphtongues nasales dans le cadre de la communauté linguistique de Québec. En 1971, Charbonneau publiait une étude très détaillée sur les voyelles nasales du français canadien. Il démontrait clairement l'existence "d'une voyelle nasale canadienne plus longue que la voyelle nasale du français général" ainsi qu'"un changement de timbre à l'intérieur de leur émission". Cette variante linguistique est corroborée par les études de Gendron (en 1966, sur les tendances phonétiques du français parlé au Canada), de même que

---

[1]Pour ce travail, diphtongues et voyelles diphtonguées auront la même signification.

[2]Je tiens à remercier Diane Vincent qui m'a suggéré de travailler sur les diphtongues et apporté une aide inestimable par ses suggestions et critiques pertinentes tout au long de cette étude. Je remercie également Denise Deshaies de m'avoir permis d'accéder au corpus Deshaies et d'avoir commenté une version préliminaire de ce texte, et Claude-Émile Rochette avec qui j'ai discuté de certains aspects phonétiques concernant l'étude pilote.

par la recherche de Brichler-Labaeye (1970) sur les voyelles nasales. Elle mentionne "une curieuse tendance personnelle du sujet canadien à allonger, d'une façon inhabituelle, certaines voyelles nasales"[3].

Le phénomène de diphtongaison a été étudié dans une perspective sociolinguistique à l'intérieur du français parlé à Montréal par Cedergren, Clermont et Côté (1981) et Santerre et Millo (1978). Ils décrivent et analysent les facteurs influençant la diphtongaison sans toutefois aborder les diphtongues nasales. Cependant, la description de contraintes linguistiques permet de comprendre le conditionnement phonologique de la diphtongue.

Plus près de nous, Yaeger (1979), dans sa recherche sur le français parlé de Montréal, et Paradis (1985), dans son étude du système vocalique de Chicoutimi et Jonquière, abordent à l'aide de l'analyse acoustique le phénomène de la diphtongaison et constatent la présence de diphtongues nasales. Yaeger souligne que l'allongement est une variante souvent associée à la diphtongaison et que l'on diphtongue les voyelles nasales dans les mêmes contextes que les voyelles orales correspondantes.

Récemment, Dumas (1987), dans *Les prononciations en français québécois* indique quatre variantes diphtonguées des quatre voyelles nasales du français québécois. Laferrière (1986), dans son mémoire sur l'étude des voyelles nasales du français montréalais analysées à l'aide d'un synthétiseur à formant, fait ressortir "qu'au cours de leur apprentissage des nasales, les Montréalais apprendraient à utiliser le trait de diphtongaison comme inhérent aux voyelles nasales".

2. Méthodologie

Cette étude se base sur les données de deux sources différentes: une étude pilote réalisée en 1987 à partir de laquelle je dégagerai quelques éléments linguistiques importants pour l'analyse des variantes et un corpus plus large où reposeront les remarques et les hypothèses que l'analyse me permettra d'émettre.

2.1 *Description de l'étude pilote: le corpus Saint-Malo*

Dans le cadre d'un projet de recherche, j'ai constitué au printemps 1987 un corpus qui m'a servi d'étude pilote pour le phénomène de la diphtongaison. Ce corpus contient quatre enregistrements d'enfants de l'école Saint-Malo dans le quartier populaire Saint-Malo dans la ville de Québec. Enregistrés à la garderie de l'école, les quatre enfants expliquent leur façon de jouer et les règlements des jeux de Monopoly et d'Échelles et Serpents.

Mon but était de discriminer le taux de diphtongaison de quatre informatrices du même âge et de milieu socio-économique semblable. Les quatre voyelles nasales, [ɑ̃], [ɛ̃], [ɔ̃], [œ̃] du français standard ont été prises en considération dans ce corpus du quartier Saint-Malo. J'y ai relevé 171 voyelles nasales. Le phénomène de diphtongaison se produit avec un taux de 12.3% pour l'ensemble des nasales analysées. La voyelle [ɛ̃] atteint un taux de diphtongaison de 57% suivie de [ɑ̃] qui en possède un de 24% et de [ɔ̃] montrant un changement de l'ordre de 19%. La voyelle [œ̃] demeure une voyelle standard dans toutes ses réalisations.

Ce corpus, malgré son petit nombre d'occurrences, m'a fourni des indices sur la pertinence de certains facteurs linguistiques. Suite à cette analyse, j'ai décidé de ne pas tenir compte d'un certain

[3] L'allongement pour moi n'est pas synonyme de diphtongaison.

nombre de voyelles nasales, compte tenu des contextes d'invariances. N'ont pas été considérés parmi les occurrences les faits suivants: la voyelle nasale [œ̃] puisqu'elle se rencontre presque uniquement dans le morphème *un* (article ou chiffre) ainsi que les morphèmes *on, en* et autres clitiques[4] qui ne se diphtonguent jamais et les variantes nasales du *Oui* (*ouin, ouan, ouain*). Le morphème *dans* suivi d'un article défini féminin comme *dans la maison* se transformant *dans'maison* a aussi été rejeté parce que l'allongement y est davantage lié à l'élision du L. Les facteurs linguistiques identifiés comme pertinents seront décrits dans la description du second corpus de cette étude.

Cette étude pilote m'a amenée à m'interroger sur le rôle du contexte social. Suite à cette réflexion, j'ai décidé d'analyser le phénomène de la diphtongaison à partir d'un échantillon sociolinguistique plus significatif extrait du corpus Deshaies de 1977 "Le français parlé dans la ville de Québec". J'observerai les conditions de réalisation des trois variantes des voyelles nasales [ɑ̃], [ɛ̃], [ɔ̃] et j'analyserai les différences d'usage de ces variantes selon l'âge des locuteurs et les groupes sociaux auxquels ils appartiennent. Je comparerai finalement les résultats obtenus pour les deux corpus.

3. Le corpus Deshaies

À partir d'environ deux heures d'enregistrements tirés du corpus Deshaies (1977), j'ai procédé à l'analyse auditive des trois voyelles nasales. À raison de trois à quatre minutes choisies au hasard pour chacun des vingt informateurs, ma tâche était de discriminer les trois variantes des nasales. Cet échantillon tient compte de trois groupes d'âge; quatre enfants âgés de dix ans, huit adolescents d'environ seize ans ainsi que huit adultes. Ils se répartissent, de façon mixte et égale, dans deux groupes sociaux: un groupe du quartier Saint-Sauveur représentant la classe populaire du centre de Québec et un autre de Sainte-Foy représentant la classe moyenne de Québec.

3.1 *Description des données*

Au total, 1011 voyelles nasales ont été considérées. Parmi celles-ci, on observe 15 diphtongues (1.5%) et 68 voyelles nasales allongeantes (6.7%). La nasale [ɛ̃] s'attribue le plus haut pourcentage des réalisations diphtonguées, soit 67%. Ensuite, la nasale [ɑ̃] apparaît avec un taux de 27%, suivie de [ɔ̃] qui possède 7% des productions. Ce taux de diphtongaison s'avère beaucoup moindre que les données recuellies au printemps 1987, tel que mentionné au point 2.1. Cet écart sera discuté dans notre analyse.

TABLEAU 1:

Taux de diphtongues

Corpus Deshaies	15 Diphtongues	[ɛ̃] :	10	67%
	.....	[ɑ̃] :	4	27%
	.....	[ɔ̃] :	1	7%
Étude pilote	21 Diphtongues	[ɛ̃] :	12	57%
(Saint-Malo)	.....	[ɔ̃] :	5	24%
	.....	[ɑ̃] :	4	19%

[4]Les clitiques ont comme caractéristique de se coller à un morphème. Ils ne reçoivent pas certaines variations de prononciation (Dumas,1977).

3.2 *Le contexte linguistique*

Comme mentionné au début, je me suis attardée au contexte linguistique des voyelles nasales. Des facteurs pris en compte lors de la première analyse, comme l'existence d'une pause avant ou après la syllabe contenant une voyelle nasale ont été rejetés parce qu'ils n'exercaient aucune influence dans les manifestations des variantes.

Les trois variables linguistiques qui ont retenu mon attention sont: A-la nature ouverte ou fermée de la syllabe, B-la position de la voyelle nasale dans le mot et la phrase ainsi que C-le phénomène d'accent. Ces variables linguistiques sont illustrées dans le tableau 2.

A- matin, cinq

B- avant, empêcher, attendez,
Bien, c'est tout, Il le dit bien.

C- demain, minceur

Dans mon échantillon, toutes les variantes diphtonguées se réalisent en syllabe ouverte. Il me faut souligner que cette tendance est contraire aux résultats d'autres études provenant d'autres communautés linguistiques qui attestent la présence de diphtongues seulement en syllabe fermée.

Les variantes diphtonguées se retrouvent autant dans les débuts que dans les fins de syntagmes (53% pour les débuts et 47% pour les fins). Aucune réalisation diphtonguée n'intervient en milieu de phrase. Les variantes allongées se retrouvent en plus grand nombre vers la fin des syntagmes (60%).

La position de la variante diphtonguée dans le mot est plus significative. Il apparait deux fois plus de diphtongues dans un contexte monosyllabique que dans un contexte polysyllabique (fin de mot). Là également, aucune variante diphtonguée et allongée ne se retrouve dans la première syllabe d'un mot.

Le phénomène de l'accentuation semble jouer un rôle dans la production des variantes non-standards: 66% des variantes diphtonguées et 60% des allongées sont accentuées. Cependant, le peu d'occurrences des voyelles diphtonguées m'empêche ici de pousser plus loin l'analyse linguistique.

3.3 *Le contexte sociolinguistique*

En analysant le tableau 3, qui présente la distribution des variantes selon les variables sociales, on peut remarquer pour la diphtongaison les résultats suivants: la production des diphtongues diminue avec l'âge. Elle se réalise à Sainte-Foy plus fréquemment chez les enfants pour s'amenuiser avec l'âge. Dans Saint-Sauveur, les diphtongues se retrouvent en plus grand nombre chez les adolescents et diminue sensiblement chez les enfants et plus encore dans le groupe d'adultes.

Si on compare la distribution de la variante diphtonguée chez les deux groupes sociaux selon l'âge (tableau 3), on observe que le groupe présentant le plus d'écart est celui des enfants: les enfants de Sainte-Foy possèdent davantage de variantes diphtonguées que les enfants de Saint-Sauveur. L'écart entre les groupes d'adolescents est moindre, les adolescents de Saint-Sauveur diphtonguant légèrement plus que ceux de Sainte-Foy. Pour les adultes, l'écart est quasiment nul.

TABLEAU 2:

Contexte linguistique

	Nombre de cas	Réalisations phonétiques Diphtongue	Allongée	Standard
Syllabe ouverte	984	15 (1,5%)	54 (5,5%)	915 (93,0%)
Syllabe fermée	27	- (0%)	14 (51,9%)	13 (48,1%)
Fin de syntagme	367	7 (1,9%)	41 (11,2%)	319 (86,9%)
Début de syntagme	429	8 (1,9%)	27 (6,3%)	394 (91,8%)
Milieu de syntagme	215	- (0%)	- (0%)	215 (100%%)
Contexte monosyllabe	513	11 (2,1%)	52 (10,1%)	450 (87,7%)
Contexte fin de mot	300	4 (1,3%)	16 (5,3%)	280 (93,3%)
Contexte début mot	198	- (0%)	- (0%)	198 (100%)
Accent	407	10 (2,5%)	41 (10%)	356 (87,5%)
Non Accent	604	5 (0,8%)	27 (4,5%)	572 (94,7%)

En ce qui concerne l'allongement, il se retrouve plus fortement, comme la diphtongaison, chez les enfants mais suivi de très près par le groupe d'adultes. Le groupe d'adolescents possède le plus petit taux d'allongement. Les groupes d'enfants et d'adultes du quartier de Sainte-Foy allongent avec un écart plus significatif comparativement aux groupes de Saint-Sauveur, tandis que les adolescents n'ont qu'1% d'écart entre les deux groupes.

Le sexe des locuteurs ne semble pas pertinent dans cette étude. En effet, les femmes diphtonguent dans la même proportion que les hommes. Ce qu'il faut considérer, c'est que la diphtongue nasale se retrouve plus fréquemment chez un groupe d'âge du quartier Sainte-Foy (classe moyenne) que chez ceux de ce même groupe dans Saint-Sauveur. Les enfants de Sainte-Foy représentent le groupe possédant le plus de réalisations de la variante diphtonguée. Ceci laisse supposer que cette variante ne serait donc pas une caractéristique du milieu populaire de Québec et que ce groupe a incorporé davantage la variante non-standard.

4. Analyse

Certaines hypothèses peuvent peut-être expliquer le petit nombre d'occurrences de la variante diphtonguée existant dans l'échantillon du corpus Deshaies (1977) et l'écart entre cet échantillon et le corpus pilote recueilli au printemps 1987.

4.1 *Méthodologie différente*

Une différence de méthodologie dans la cueillette des données a pu provoquer les écarts de résultats entre les deux corpus. Le corpus de 1977 provient d'entrevues plus formelles effectuées au domicile du locuteur tandis que le corpus de 1987 fut relevé en milieu scolaire (garderie de l'école) et comporte des entrevues à caractère plus informel.

TABLEAU 3:

Résultats des données

	Enfants		Adolescents		Adultes	
	D	A	D	A	D	A
Ste-Foy	4	10	3	10	2	27
	4,7%	11,6%	1,3%	4,2%	0,7%	10%
	(86 occ.)		(237 occ.)		(269 occ.)	
St-Sauv.	1	4	4	9	1	9
	1,4%	5,4%	2,4%	5,4%	0,6%	5,1%
	(74 occ.)		(167 occ.)		(178 occ.)	
Total	3,1%	8,8%	1,7%	4,7%	0,7%	8%
	(160 occ.)		(404 occ.)		(447 occ.)	

(D = variante diphtonguée A = variante allongée)

4.2 *Contexte linguistique différent*

Les variantes diphtonguées relevées dans chaque corpus présentent un contexte linguistique dissemblable. Dans le corpus de Deshaies (1977), la variante diphtonguée se réalise uniquement en syllabe ouverte et plus de la moitié des diphtongues sont accentuées (66%). Il inverse le rang des productions des variantes diphtonguées dans les mêmes proportions pour les voyelles [ɑ̃] et [ɔ̃] (tableau 1). Le corpus de Saint-Malo possède pour sa part des réalisations de variantes diphtonguées seulement sous l'accent et en position entravée. En résumé, il apparait que le phénomène d'accentuation est un facteur important quant à la fréquence de la variante dans les deux corpus.

4.3 *Le changement linguistique*

La différence des résultats obtenus pour les deux corpus amène à poser l'hypothèse du changement linguistique. Il ne faut pas oublier que je compare des résultats datant de 1977 pour le corpus Deshaies et de 1987 pour le corpus de Saint-Malo. L'écart de 10 années entre les deux échantillons peut laisser paraître un changement dans le système phonétique.

Le phénomène de diphtongaison des voyelles nasales n'est sûrement pas récent mais son insertion dans la communauté linguistique québécoise est manifeste si on compare le comportement des enfants de quartiers similaires St-Malo \ St-Sauveur à dix ans d'intervalle dans le tableau 4.

Dans l'échantillon 87, on retrouve plus de diphtongues que de variantes allongées tandis que dans le corpus de 77, il y a plus d'allongées que de diphtongues pour des enfants du même âge et du même groupe social. Cependant, même si les enfants du corpus de 77 produisent moins de variantes diphtonguées que les enfants de l'étude pilote, ils en produisent dans des contextes linguistiques différents et plus que les autres groupes d'âge du même corpus. Avec des données plus considérables, je pourrais sérieusement envisager un début de changement linguistique dont les enfants seraient les innovateurs.

TABLEAU 4:

Diphtongues et allongeantes chez des enfants de quartiers populaires

Corpus 87 (St-Malo)		Corpus 77 (St-Sauveur)		Corpus 77 (S.F)	
D	A	D	A	D	A
21	12	1	4	4	10
12,3%	7,0%	1,4%	5,4%	4,7%	11,6%
(171 occ.)		(74 occ.)		(86 occ.)	

4.4 *L'origine et diffusion géographique de la variante*

Je peux formuler une dernère hypothèse, celle de l'origine de la variante et de sa diffusion géographique. Mon intuition et la discussion avec des linguistes québécois me porte à croire que la diphtongaison des voyelles nasales, très fréquente dans certaines banlieues de Québec, a pu s'infiltrer peu à peu à Québec, compte tenu du déplacement de la population des banlieues vers la ville pour le travail. Cette hypothèse expliquerait, en partie, la saillance de la variante. En effet, on peut difficilement croire qu'une variante produite dans 1% des cas devienne un stigmate d'une communauté linguistique. La production de diphtongues nasales serait en fait une caractéristique des gens de l'extérieur de la communauté linguistique de Québec.

Il est important de remarquer le peu de différences entre les résultats des quartiers Sainte-Foy et Saint-Sauveur. Ceci peut indiquer un rétrécissement d'écart entre le français parlé de ces deux quartiers de Québec depuis quelques années. Le quartier de Sainte-Foy ne serait plus l'antithèse de celui de Saint-Malo et Saint-Sauveur.

5. Conclusion

L'origine des variantes non standard et leur diffusion à l'intérieur de la ville de Québec est un phénomène intéressant qui s'étudie en tenant compte de la sociolinguistique et de la géolinguistique. Cette étude est, en quelque sorte, une tentative de débroussaillement du phénomène des diphtongues dans la communauté linguistique de Québec. J'ai tenté de montrer tout au long de cette analyse que l'apparition des variantes diphtonguées est favorisée par des facteurs linguistiques (accent, contexte, etc.) et sociaux (âge, groupes sociaux) et qu'elles correspondent à un changement récent dans la communauté, véhiculé principalement par les plus jeunes.

Je ne prétends pas avec cette analyse répondre à toutes les questions mais elle servira au moins à démontrer que toutes les hypothèses formulées et les éventuelles études n'auraient pour objectif que d'expliquer la saillance de la diphtongue nasale : comment en effet une variante non standard a-t-elle pu devenir un stigmate du parler de la ville de Québec alors qu'elle en est presque absente? Cette situation indique peut-être qu'une basse fréquence d'un phénomène est suffisante à l'élaboration de stéréotypes.

BIBLIOGRAPHIE

BRICHLER-LABAEYE, C. (1970), *Les voyelles françaises, mouvements et positions articulatoires à la lumière de la radiocinématographie*, Paris, Klincksieck, 258 p.

CEDERGREN, H., CLERMONT, J. et COTÉ, F. (1981), "Le facteur temps et deux diphtongues du français montréalais", dans D. Sankoff et H. Cedergren (éds), *Variation Omnibus*, Edmonton, Canada, pp.169-176.

CHARBONNEAU, René (1971), *Étude sur les voyelles nasales du français canadien*, Québec - Paris: Presses de l'Université Laval - Librairie C. Klincksieck.

DUMAS, Denis (1974), "Durée vocalique et diphtongaison en français québécois", *Le français de la région de Montréal. Aspects phonétique et phonologique*, Montréal, Presses de l'Université du Québec.

DUMAS, Denis (1981), "Structure de la diphtongaison québécoise", dans *Canadian Journal of Linguistics / Revue canadienne de linguistique*, vol. 26, no.1, pp. 1-61.

DUMAS, Denis (1987), *Les prononciations en français québécois. Nos façons de parler*, Québec, Presses de l'Université du Québec. 150 p.

GENDRON, Jean-Denis (1966), *Tendances phonétiques du français parlé au Canada*, Paris-Québec: Klincksieck - Presses de l'Université Laval.

LABOV, William (1972), *Sociolinguistic patterns*, Philadelphia, University of Pennsylvania Press.

LAFERRIÈRE, François (1986), *Analyse synthèse et étude de règles acoustiques de production avec un synthétiseur à formants*, Mémoire, Université du Québec.

PARADIS, Claude (1985), *An acoustic study of variation and change in the vowel system of Chicoutimi and Jonquiere (Quebec)*, Unpublished Ph.D. dissertation, University of Pennsylvania.

SANTERRE, L. et MILLO, J. (1978), "Diphthongization in Montreal" dans D. Sankoff (éd), *Linguistic Variation*, Université de Montréal, Montréal, pp. 173-184.

YAEGER, Malcah (1979), *Context-determined variation in Montreal French vowels*, Unpublished Ph.D. dissertation, University of Pennsylvania.